

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 14, 2004, 15:03:35 ; Search time 13362 Seconds  
(without alignments)  
6967.582 Million cell updates/sec  
Title: US-09-813-329-5  
Perfect score: 2148  
Sequence: 1 ggcacggcgacgacgt.....aaaaaaaaaactcgag 2148  
Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 100%  
Listing first 45 summaries

Database :			GenEmbl.*							
1:	gb	ba.*	22	59.6	2.8	229583	10	AL844566	Mouse DNA	
2:	gb	htg.*	C	23	59.4	2.8	259427	2	AC137318	Rattus no
3:	gb	in.*	C	24	59.4	2.8	284938	2	AC091335	Rattus no
4:	gb	om.*	C	25	58.8	2.7	643	3	AY301008	Pogonomyr
5:	gb	ov.*	C	26	58.6	2.7	234817	10	AL663048	Mouse DNA
6:	gb	pat.*	C	27	58	2.7	3127	14	AF305694	Kaposi's a
7:	gb	ph.*	C	28	56.8	2.6	206973	2	AC130594	Rattus no
8:	gb	pl.*	C	29	56.6	2.6	247900	2	AC136408	Rattus no
9:	gb	pr.*	C	30	56.2	2.6	223508	2	AC102480	Mus muscu
10:	gb	ro.*	C	31	55.8	2.6	3489	6	AR072952	Sequence
11:	gb	sts.*	C	32	55.8	2.6	3489	6	AR254713	Sequence
12:	gb	ty.*	C	33	55.8	2.6	3489	6	AR261360	Sequence
13:	gb	un.*	C	34	55.8	2.6	3489	6	AX107938	Sequence
14:	gb	vi.*	C	35	55.8	2.6	3489	14	KSU52064	Kaposi's a
15:	em	ba.*	C	36	55.8	2.6	32207	6	AR065852	Sequence
16:	em	fun.*	C	37	55.8	2.6	32207	6	AR127850	Sequence
17:	em	hum.*	C	38	55.8	2.6	32207	6	AR194752	Sequence
18:	em	in.*	C	39	55.8	2.6	32207	6	BD190717	Unique as
19:	em	mu.*	C	40	55.8	2.6	137508	14	KSU75698	U75698 Kaposi's a
20:	em	om.*	C	41	55.6	2.6	620	9	BC035567	Homo sapi
21:	em	or.*	C	42	55.6	2.6	1847	3	AY069729	Drosophil
22:	em	ov.*	C	43	55.6	2.6	2931	14	AF192756	Kaposi's a
23:	em	pat.*	C	44	55.4	2.6	1779	3	AB104617	Epithelium
			C	45	55	2.6	88549	3	AC116979	Dictyoste

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2103.8	97.9	2165	3	BT001838	BT001838 Drosophil
2	2101	97.8	2101	3	AF149799	AF149799 Drosophil
3	2076.4	96.7	2159	3	AF521176	AF521176 Drosophil
4	1656	77.1	1656	3	AB073865	AB073865 Drosophil
5	1202	56.0	1248	3	AY115551	AY115551 Drosophil
6	1149.4	53.5	1221	3	AY119233	AY119233 Drosophil
7	798.8	37.2	55359	2	AC012854	AC012854 Drosophil
8	798.8	37.2	158983	3	AC005974	AC005974 Drosophil
9	798.8	37.2	172904	3	AC007414	AC007414 Drosophil
10	798.8	37.2	189620	3	AC099029	AC099029 Drosophil
11	798.8	37.2	275390	3	AE003831	AE003831 Drosophil
12	67.2	3.1	2040	5	AB117518	AB117518 Theragra
13	67.2	3.1	2235	5	AB117517	AB117517 Theragra
14	66.8	3.1	390	6	AR135147	AR135147 Sequence
15	66.4	3.1	2000	6	AX655393	AX655393 Sequence
16	65.2	3.0	148198	2	BX276082	BX276082 Danio rer
17	64.4	3.0	215119	2	AC098082	AC098082 Rattus no
18	62.4	2.9	143586	5	AL627248	AL627248 Zebrafish
19	62.4	2.9	148339	2	BX677670	BX677670 Danio rer
20	60.8	2.8	232092	2	AC127800	AC127800 Rattus no
21	60.4	2.8	243597	2	AC099373	AC099373 Rattus no
22	59.6	2.8	229583	10	AL844566	AL844566 Mouse DNA
23	59.4	2.8	259427	2	AC137318	AC137318 Rattus no
24	59.4	2.8	284938	2	AC091335	AC091335 Rattus no
25	58.8	2.7	643	3	AY301008	AY301008 Pogonomyr
26	58.6	2.7	234817	10	AL663048	AL663048 Mouse DNA
27	58	2.7	3127	14	AF305694	AF305694 Kaposi's
28	56.8	2.6	206973	2	AC130594	AC130594 Rattus no
29	56.8	2.6	247900	2	AC136408	AC136408 Rattus no
30	56.2	2.6	223508	2	AC102480	AC102480 Mus muscu
31	55.8	2.6	3489	6	AR072952	AR072952 Sequence
32	55.8	2.6	3489	6	AR254713	AR254713 Sequence
33	55.8	2.6	3489	6	AR261360	AR261360 Sequence
34	55.8	2.6	3489	6	AX107938	AX107938 Sequence
35	55.8	2.6	3489	14	KSU52064	US2064 Kaposi's sa
36	55.8	2.6	32207	6	AR065852	AR065852 Sequence
37	55.8	2.6	32207	6	AR127850	AR127850 Sequence
38	55.8	2.6	32207	6	AR194752	AR194752 Sequence
39	55.8	2.6	32207	6	BD190717	BD190717 Unique as
40	55.8	2.6	137508	14	KSU75698	U75698 Kaposi's sa
41	55.6	2.6	620	9	BC035567	BC035567 Homo sapi
42	55.6	2.6	1847	3	AY069729	AY069729 Drosophil
43	55.6	2.6	2931	14	AF192756	AF192756 Kaposi's
44	55.4	2.6	1779	3	AB104617	AB104617 Epidinium
45	55	2.6	88549	3	AC116979	AC116979 Dictyoate

ALIGNMENTS

RESULT 1	BT001838	BT001838	2165 bp	mRNA	linear	INV 15-NOV-2002
LOCUS	Drosophila	melanogaster	RH51659	full	insect	cdna.
DEFINITION	BT001838					
ACCESSION	BT001838					
VERSION	BT001838.1	GI:25013017				
KEYWORDS	FLI_CDNA.					
SOURCE	Drosophila melanogaster (fruit fly)					
ORGANISM	Drosophila melanogaster					
	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;					
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;					
	Ephydroidea; Drosophilidae; Drosophila.					
REFERENCE	1 (bases 1 to 2165)					
AUTHORS	Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,					
	Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,					

George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G., Miranda, A., Mungall, C. J., Nunco, J., Pacleb, J., Paragas, V., Park, S., Patel, S., Phuanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celisner, S.

Direct Submission  
Submitted (15-NOV-2002) Berkeley Drosophila Genome Project,  
Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, USA

Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to be reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to [cdna@fruitfly.berkeley.edu](mailto:cdna@fruitfly.berkeley.edu).

FEATURES  
source Location/Qualifiers  
1..2165  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/strain="Y; cn bw sp"  
/db\_xref="taxon:7227"

misc\_feature 1..2152  
/note="sim4 alignment with AE003831 (46B13-46E4)"  
648..1877

CDS  
/note="Longest ORF"  
/codon\_start=1  
/product="RH51659p"  
/protein\_id="AAN71595.1"  
/db\_xref="GI:25013018"

/translation="MTAEILKPFITPSANDGPPAKATATAQRTROLIPLVLGFIGLVVAIIALITQWTRVYSLHDKLSKRVVDNLQRLGNYLDEDFEFOKEYENALIDPKKVDLTDDEDDDDGDLSDADDDDDVSVDVGDVDEYDMLNKLNAHGTPTSETTAEGGETDSASANDNDVDFDFTSYNAHKKOKERSRSTADVRNEONIQGNHTELOKSNEATSKERPAPLHRRMRSHRHLVVRKARSDSRPAHFHLSSRRHQSMGHGDMYIGNERNSTOCHFTQRDVLTVTNIGLYIYVIAQICNNSHDQNGFIVFGDDTFFLOCLNTVPTNMPHKVHTCHTSLHLERNERIHLKDIHNDNRAVLREGNNRSYFGIFKV"

ORIGIN

Query Match 97.9%; Score 2103.8; DB 3; Length 2165;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2105; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 5 CGAGCGGAAACGAGCGTTTAAAGTGAAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64  
Db 19 CGACGCGAACGAGCGTTTAAAGTGAAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 78  
  
Qy 65 CGCGTGGTGCATGATAGCGAAGAAAAAGCTATCCGTTTCAGTTTAACTTACCTTACCAAG 124  
Db 79 CGCGTGGTGCATGATAGCGAAGAAAAAGCTATCCGTTTAACTTACCTTACCAAG 138  
  
Qy 125 ATTGAATTTGCCCATCGGCAAAATTTACTAAATAATACATAAGTGCACCTCGCTCCACTGTGT 184  
Db 139 ATTGAATTTGCCCATCGGCAAAATTTACTAAATAATACATAAGTGCACCTCGCTCCACTGTGT 198  
  
Qy 185 GTTGTGTTTTTTTTTTTTTTTTTGGTTTTTCGCTGTGCTTTATCGCAAAACAAGAACTGAT 244  
Db 199 GTTGTGTTTTTTTTTTTTTTTTTGGTTTTTCGCTGTGCTTTATCGCAAAACAAGAACTGAT 258  
  
Qy 245 AAAACTAGAAAATATCTTGAGAAACTTGTTCGCGCTTTCTTTTGTCTTAATTTGCCGATC 304  
Db 1339 AGGAAAAGTGTCCAAATGAGGCAACTTCCAAAGAGAGACCTTGCACCTTCAACCCCGTC 1398

Db 259 AAACTAGAAAATATCTTGAGAAACTTGTTCGCGCTTTCTTTTGTCTTAATTTGCCGATC 318  
Qy 305 GCGGAGAGAAAAACAAGCAAGTGAAGCAAAACAAGTGTGGTAAATACAAATCTGAAAAAGGCA 364  
Db 319 GCGGAGAGAAAAACAAGCAAGTGAAGCAAAACAAGTGTGGTAAATACAAATCTGAAAAAGGCA 378  
Qy 365 CCATCAGCAGCCCGAGGGTTTATCTATATAGATGTCGAGCTTATCATCTCATCTGCTGTC 424  
Db 379 CCATCAGCAGCCCGAGGGTTTATCTATATAGATGTCGAGCTTATCATCTCATCTGCTGTC 438  
Qy 425 TGTGAGGTGTTCTTCTGCTGCTAGTATCTTAAATACATAGAGTGTGTTTATATATAA 484  
Db 439 TGTGAGGTGTTCTTCTGCTGCTAGTATCTTAAATACATAGAGTGTGTTTATATATAA 498  
Qy 485 GTCCGACAAAGCTCGATTGGAACACAGTGTGAGTGCCTTTGAGTGGGTGGGCAAGATCG 544  
Db 499 GTCCGACAAAGCTCGATTGGAACACAGTGTGAGTGCCTTTGAGTGGGTGGGCAAGATCG 558  
Qy 545 TCATCATCATCATCTGCTGCTCATTCACAGAAATCAGCATCAGCATCTCGAGGCCCGGT 604  
Db 559 TCATCATCATCATCTGCTGCTCATTCACAGAAATCAGCATCAGCATCTCGAGGCCCGGT 618  
Qy 605 TGCTCTAAGATCCCAAGTGTTCATCAATATGATCGCCGAGACCCCTCAAGCCGTTTATAA 664  
Db 619 TGCTCTAAGATCCCAAGTGTTCATCAATATGATCGCCGAGACCCCTCAAGCCGTTTATAA 678  
Qy 665 CGCCAAAGAGTGCACAGATGATGTTTCCGGCCAAAGCGACGACGCGGACCGCC 724  
Db 679 CGCCAAAGAGTGCACAGATGATGTTTCCGGCCAAAGCGACGACGCGGACCGCC 738  
Qy 725 AGCGACGACCCCGCAGCTGATCCCTCGTGTGTTTGGGGTTCATCGGTCTGGGGCTGGTCG 784  
Db 739 AGCGACGACCCCGCAGCTGATCCCTCGTGTGTTTGGGGTTCATCGGTCTGGGGCTGGTCG 798  
Qy 785 TTGCCATTCTCGACCTAACGATCTGSCAGACAAACGCTGTATCGCATCTGGCAAGAGAC 844  
Db 799 TTGCCATTCTCGACCTAACGATCTGSCAGACAAACGCTGTATCGCATCTGGCAAGAGAC 858  
Qy 845 TGAAGAGCTGAAGCGAGTCTGATGATCTCCAGCAGCGTTTGGGCATAAATCTATCTGG 904  
Db 859 TGAAGAGCTGAAGCGAGTCTGATGATCTCCAGCAGCGTTTGGGCATAAATCTATCTGG 918  
Qy 905 AGAGTTTCAGCAGATTCCAAAAGAGTACGAAATGCCCTCATCGACTATCAAAAAAGG 964  
Db 919 AGAGTTTCAGCAGATTCCAAAAGAGTACGAAATGCCCTCATCGACTATCCAAAAAGG 978  
Qy 965 TGGATGGCTTCAAGATGAGGAGACGACGATGGCGATGGTCTGGAATCCATTGGCG 1024  
Db 979 TGGATGGCTTCAAGATGAGGAGACGACGATGGCGATGGTCTGGAATCCATTGGCG 1038  
Qy 1025 AGCAGGAGACGACGATTTAGCTATAGCTCTGATGATGATGTTGGCGGAGACTAGAGG 1084  
Db 1039 AGCAGGAGACGACGATTTAGCTATAGCTCTGATGATGATGTTGGCGGAGACTAGAGG 1098  
Qy 1085 ACTTACACCGATATGTTAAATAAATCAACATGCAATACCGGACCGCCACCATCTG 1144  
Db 1099 ACTTACACCGATATGTTAAATAAATCAACATGCAATACCGGACCGCCACCATCTG 1158  
Qy 1145 AGACCACTGCTGAGGGCGAGGCGAGACGACAGTGCATCTCAGCTCAAAATGATGACA 1204  
Db 1159 AGACCACTGCTGAGGGCGAGGCGAGACGACAGTGCATCTCAGCTCAAAATGATGACA 1218  
Qy 1205 ATGTGTTGATGACTTTTACCAGCTACATGCCCAACAAAAGAGAGAGAGAAATCTC 1264  
Db 1219 ATGTGTTGATGACTTTTACCAGCTACATGCCCAACAAAAGAGAGAGAGAAATCTC 1278  
Qy 1265 GCTCGATTGCCGATGACCAATGAGGAGCAATATTTCAAGGAAATCACACAGAGCTTC 1324  
Db 1279 GCTCGATTGCCGATGACCAATGAGGAGCAATATTTCAAGGAAATCACACAGAGCTTC 1338  
Qy 1325 AGGAAAAGTGTCCAAATGAGGCAACTTCCAAAGAGAGCGCTGCACACTTACACCGCTC 1384  
Db 1339 AGGAAAAGTGTCCAAATGAGGCAACTTCCAAAGAGAGCGCTGCACACTTACACCGCTC 1398



QY 610 TAAGATCCCGAGTGTTCATCAATTATGACTGCCGAGACCCCTCAAGCCGTTTATTAACGCCA 669  
DB 601 TAAGATCCCGAGTGTTCATCAATTATGACTGCCGAGACCCCTCAAGCCGTTTATTAACGCCA 660  
QY 670 ACGAGTCCCAACGATGATGGTGTTCGCGCCAAAGCGACGACGACGCGCGCCAGCGA 729  
DB 661 ACGAGTCCCAACGATGATGGTGTTCGCGCCAAAGCGACGACGACGCGCGCCAGCGA 720  
QY 730 CGCACCCGCGAGCTGATCCCCCTGGTTTGGGTTTCATCGGTCTGGGCTGGTGGTGGC 789  
DB 721 CGCACCCGCGAGCTGATCCCCCTGGTTTGGGTTTCATCGGTCTGGGCTGGTGGTGGC 780  
QY 790 ATTCTCGCACTAAACGATCTGCGACACGCGGTGTATCGCATCTGGACAAGGACTGAAG 849  
DB 781 ATTCTCGCACTAAACGATCTGCGACACGCGGTGTATCGCATCTGGACAAGGACTGAAG 840  
QY 850 AGCCTGAAGCGAGTCTGCTGATAATCTCCAGCAGCGGTTTGGGCATAAACTATCTGGACGAG 909  
DB 841 AGCCTGAAGCGAGTCTGCTGATAATCTCCAGCAGCGTGTGGGCATAAACTATCTGGACGAG 900  
QY 910 TTCGACGAGTTCBAAAAGGAGTACGAGAAATGCCCTCATCTGACTATCCAAAAAAGGTGGAT 969  
DB 901 TTCGACGAGTTCBAAAAGGAGTACGAGAAATGCCCTCATCTGACTATCCAAAAAAGGTGGAT 960  
QY 970 GGCCTCAGGATGAGGAGGACGACGATGGCGATGCTGTGGATTCATTTGGCGGACGAC 1029  
DB 961 GGCCTCAGGATGAGGAGGACGACGATGGCGATGCTGTGGATTCATTTGGCGGACGAC 1020  
QY 1030 GAGGACGACGAGTTAGCTATAGTCTGTGGATGATGTTGGCGCAGACTACGAGGACTAC 1089  
DB 1021 GAGGACGACGAGTTAGCTATAGTCTGTGGATGATGTTGGCGCAGACTACGAGGACTAC 1080  
QY 1090 ACCGATATGTTAAATAAATCAACATGCGACATACCGGACCGACCGCCACATCTGAGACC 1149  
DB 1081 ACCGATATGTTAAATAAATCAACATGCGACATACCGGACCGACCGCCACATCTGAGACC 1140  
QY 1150 ACTGCTGAGGCGAGGCGGACGACGATGCTCATCTCAGCTCAATGATGACAAATGTCG 1209  
DB 1141 ACTGCTGAGGCGAGGCGGACGACGATGCTCATCTCAGCTCAATGATGACAAATGTCG 1200  
QY 1210 TTCGATGACTTTACGAGTACAAATGCCACAAAAGAGCAGAGAGAAAATCTCGCTCG 1269  
DB 1201 TTCGATGACTTTACGAGTACAAATGCCACAAAAGAGCAGAGAGAAAATCTCGCTCG 1260  
QY 1270 ATTCCCGATGACCAATGAGGACGAGAAATTTCAAGGAATACACAGAGCTTCAGGAA 1329  
DB 1261 ATTCCCGATGACCAATGAGGACGAGAAATTTCAAGGAATACACAGAGCTTCAGGAA 1320  
QY 1330 AAGTCATCCATGAGGCAACTTCCAAAGAGAGCCCTGCACCACTTCACCAACCGTCGCGA 1389  
DB 1321 AAGTCATCCATGAGGCAACTTCCAAAGAGAGCCCTGCACCACTTCACCAACCGTCGCGA 1380  
QY 1390 ATGCATTTCCGCGCATCGCCACCTCTAGTCGCGAAAGCCAGATCCGAGGACTCGAGGCCA 1449  
DB 1381 ATGCATTTCCGCGCATCGCCACCTCTAGTCGCGAAAGCCAGATCCGAGGACTCGAGGCCA 1440  
QY 1450 GCAGCCCATTTCCATTTGAGCAGCAGCGGCGGTACCACGAAGATATGGCTACCATGGA 1509  
DB 1441 GCAGCCCATTTCCATTTGAGCAGCAGCGGCGGTACCACGAAGATATGGCTACCATGGA 1500  
QY 1510 GATATGTACATAGGAATGATTAACGAGAGAACTCTTATCAGGACACATTTTCAAAACGCGC 1569  
DB 1501 GATATGTACATAGGAATGATTAACGAGAGAACTCTTATCAGGACACATTTTCAAAACGCGC 1560  
QY 1570 GATGGCTCTTGACGCTGACCAATACAGGCTATATTAACGATACGCGCCAGATATGCTAC 1629  
DB 1561 GATGGCTCTTGACGCTGACCAATACAGGCTATATTAACGATACGCGCCAGATATGCTAC 1620  
QY 1630 AACAACTCGCAGCAGACGAGAGGATTTATCGCTTTTCAAGGACGACATCCATTTCTGAC 1689  
DB 1621 AACAACTCGCAGCAGACGAGAGGATTTATCGCTTTTCAAGGACGACATCCATTTCTGAC 1680

QY 1690 TGCTTGAACACGGTGCCCAACCATGCCACATTAAGTGCCACACCTGCCACACGAGTGGT 1749  
DB 1681 TGCTTGAACACGGTGCCCAACCATGCCACATTAAGTGCCACACCTGCCACACGAGTGGT 1740  
QY 1750 CTGATCCACCTCGAAACGAAACGAGAGATCCATCTGAAGGACATTCACACGATCGCAAT 1809  
DB 1741 CTGATCCACCTCGAAACGAAACGAGAGATCCATCTGAAGGACATTCACACGATCGCAAT 1800  
QY 1810 GCAGTTCTCGCGGAGGAAACAAACGAAAGCTACTTTTGGCATCTTCAAGGTGTAATTTGGA 1869  
DB 1801 GCAGTTCTCGCGGAGGAAACAAACGAAAGCTACTTTTGGCATCTTCAAGGTGTAATTTGGA 1860  
QY 1870 GAGATTATCCCGGTGAGAGATGGAATACCACTTTAAGCTTTTGTCCCGGACCTGCTC 1929  
DB 1861 GAGATTATCCCGGTGAGAGATGGAATACCACTTTAAGCTTTTGTCCCGGACCTGCTC 1920  
QY 1930 GTGAATGCCGATTCATCGCCAGCGTGAAATCCATTAGTTCGTAGTACCTTAGTCTTAGTCACT 1989  
DB 1921 GTGAATGCCGATTCATCGCCAGCGTGAAATCCATTAGTTCGTAGTACCTTAGTCTTAGTCACT 1980  
QY 1990 CCAAACTTAATCTCAATCGGAATCGTGCACTATGCTAGTTCAGAGACGAGGAAAAATC 2049  
DB 1981 CCAAACTTAATCTCAATCGGAATCGTGCACTATGCTAGTTCAGAGACGAGGAAAAATC 2040  
QY 2050 ATATTATTTTGTATATCTCGTTCCGACTCTAAAGAGTGAATAAATAATATATAGTACTAT 2109  
DB 2041 ATATTATTTTGTATATCTCGTTCCGACTCTAAAGAGTGAATAAATAATATATAGTACTAT 2100  
QY 2110 T 2110  
DB 2101 T 2101

RESULT 3  
AF521176  
LOCUS  
DEFINITION  
AF521176  
ACCESSION  
AF521176.1  
GI:21717645  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 2159)  
Moreno,E., Yan,M. and Basler,K.  
Evolution of TNF Signaling Mechanisms. JNK-Dependent Apoptosis  
Triggered by Eiger, the Drosophila Homolog of the TNF Superfamily  
Curr. Biol. 12 (14), 1263-1268 (2002)  
12176339  
2 (bases 1 to 2159)  
Moreno,E., Yan,M. and Basler,K.  
Direct Submission  
Submitted (13-JUN-2002) Institut Molecular Biology, Uni Zurich,  
Winterthurerstrasse 190, Zurich, Z 8057, Switzerland  
Location/Qualifiers  
1. .2159  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
641..1888  
/codon\_start=1  
/product="eiger"  
/protein\_id="AAM76710.1"  
/db\_xref="GI:21717646"  
CDS  
FEATURES  
source  
IGLGLVAILALTITWQTRVSHLDKELSLKRVNLOQLGINVLDSEDFQKEVFN  
ALIDYPKVDGLTDEEDDDGLDLSIADDEDDVYSVSDVDGADYEDYTDMLKLN  
NAHTGTTTSETTAEGETDSASASNDNVDFDFTSINAHKKQERKSRSIADVRN  
BEQNTQGNHTLBEQKSNEATSKESPAHLHRRHMRSHRLLHLLKQERKSLSAEDE  
RPAAFHLSRRRHQSGVGHFVIGNDNERNYSYQGFOTRDLGVTITVTNGLYVYA  
QTCYNNSHDQNGFIVFGDTPFLQCLNTVPTNMPHKVHTCHTSGLIHLERNRIHLKD



1HNDRNAVLREGNNRSYFQPKV"									
ORIGIN									
Query Match 96.7%; Score 2076.4; DB 3; Length 2159;									
Best Local Similarity 98.7%; Pred. No. 0;									
Matches 2121; Conservative 0; Mismatches 6; Indels 21; Gaps 2;									
Qy	5	CGAGCGCAACGGACGCTTTAAAGTGAAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG	64						
Db	15	CGAAGCGAACGGACGCTTTAAAGTGAAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG	74						
Qy	65	CGCGTGCCTGCATGTACGGAAGAAAAAGCTATCCGTTTCAGTTAACTACTTACCAAG	124						
Db	75	CGCGTGCCTGCATGTACGGAAGAAAAAGCTATCCGTTTCAGTTAACTACTTACCAAG	134						
Qy	125	ATTGAATTTTCGCGCATCGGGCAATTAATAAATAACATAAGTGCACCTCGTCCACTGTGT	184						
Db	135	ATCGAATTTTCGCGCATCGGGCAATTAATAAATAACATAAGTGCACCTCGTCCACTGTGT	194						
Qy	185	GTGTGTGTTTTTTTTTTTTTTTTTGGTTTTTTCGCTGTGCGCTTTTATCGCAACCAAGAACTGAT	244						
Db	195	GTGTGTG---TTTTTTTTTTTTTTTTTGGTTTTTTCGCTGTGCGCTTTTATCGCAACCAAGAACTGAT	251						
Qy	245	AAAACCTAGAAAAATATCTTTAGAAAACTTGTTTTTCGCGCTTTTCTTTTGTGCTAATTTGCCGATC	304						
Db	252	AAAACCTAGAAAAATATCTTTAGAAAACTTGTTTTTCGCGCTTTTCTTTTGTGCTAATTTGCCGATC	311						
Qy	305	CGGGAAGAGAAAAACAACGAGTAGACAAAAACAAGTGTGTAATAACAATCTGAAAGGGCA	364						
Db	312	CGGGAAGAGAAAAACAACGAGTAGACAAAAACAAGTGTGTAATAACAATCTGAAAGGGCA	371						
Qy	365	CCATCAGCAGCCCGAGGGGTTTTATCTATATAGATGTCGAGCTTATCATCTCATCTGTC	424						
Db	372	CCATCAGCAGCCCGAGGGGTTTTATCTATATAGATGTCGAGCTTATCATCTCATCTGTC	431						
Qy	425	TGTGAGGTGTTCTGTGTGCTCGTGTATCTTTAAATACATAGAGTGTGTTTATATAA	484						
Db	432	TGTGAGGTGTTCTGTGTGCTCGTGTATCTTTAAATACATAGAGTGTGTTTATATAA	491						
Qy	485	GTGGACAAAGCTCGATTGGAAACAGCTGTCGAGTGCCCTTTCAGTGGGTGGCGAGATCG	544						
Db	492	GTGGACAAAGCTCGATTGGAAACAGCTGTCGAGTGCCCTTTCAGTGGGTGGCGAGATCG	551						
Qy	545	TCATCATCATCATCTCGTCGCTATTACAAAGAAATCAGCATCAGCATCTGGAGGCCCGGT	604						
Db	552	TCATCATCATCATCTCGTCGCTATTACAAAGAAATCAGCATCAGCATCTGGAGGCCCGGT	611						
Qy	605	TGCTCTAAGATCCCGAGTTCATCAATATATGACTGCGGAGACCTCAAGCCGTTTATAA	664						
Db	612	TGCTCTAAGATCCCGAGTTCATCAATATATGACTGCGGAGACCTCAAGCCGTTTATAA	671						
Qy	665	CGCCAAACAGTGCACCATGATGTTTTTCGGCCAAAGCAGCAGCGGACCGCCC	724						
Db	672	CGCCAAACAGTGCACCATGATGTTTTTCGGCCAAAGCAGCAGCGGACCGCCC	731						
Qy	725	AGCGACGACCCCGCAGCTGATCCCGCTGGTTTTTGGGGTTTCATCGGTCTGGGGTGGTCG	784						
Db	732	AGCGACGACCCCGCAGCTGATCCCGCTGGTTTTTGGGGTTTCATCGGTCTGGGGTGGTCG	791						
Qy	785	TTGCCATTTCTGCATTAACGATCTGGCAGACAAACCGGTGTATCGCATCTGGACAAAGGAGC	844						
Db	792	TTGCCATTTCTGCATTAACGATCTGGCAGACAAACCGGTGTATCGCATCTGGACAAAGGAGC	851						
Qy	845	TGAAGAGCTGAGGAGCTCGTCGATATCTCCAGCAGCGTTTGGGCATAAATCTCTGG	904						
Db	852	TGAAGAGCTGAGGAGCTCGTCGATATCTCCAGCAGCGTTTGGGCATAAATCTCTGG	911						
Qy	905	ACGAGTTCGACGAGTTCCTAAAGAGGATACGAGAAATGCCCTCATCGACTATCCAAAAAAGG	964						
Db	912	ACGAGTTCGACGAGTTCCTAAAGAGGATACGAGAAATGCCCTCATCGACTATCCAAAAAAGG	971						
Qy	965	TGGATGGCCTCACCGATGAGGAGGACGACGATGGCGATGGTCTGGATTTCCATTTGCCG	1024						
Db									

Db	972	TGGATGGCCTCACCGATGAGGAGGACGACGATGGCGATGGTCTGGATTCATTTGCGG	1031
Qy	1025	ACGACGAGGACGACGAGCTTAGCTATAGCTCTGTGGATGATGTTGGCGCAGACTACGAGG	1084
Db	1032	ACGACGAGGACGACGAGCTTAGCTATAGCTCTGTGGATGATGTTGGCGCAGACTACGAGG	1091
Qy	1085	ACTACACCGATATGTTAAATAAACTCAACAATGACATACCGGCACCAACGCCCATCTG	1144
Db	1092	ACTACACCGATATGTTAAATAAACTCAACAATGACATACCGGCACCAACGCCCATCTG	1151
Qy	1145	AGACCACTGCTGAGGGCGAGGGCGAGACGGACAGTGCATCTCTCAGCTCAATATGATGACA	1204
Db	1152	AGACCACTGCTGAGGGCGAGGGCGAGACGGACAGTGCATCTCTCAGCTCAATATGATGACA	1211
Qy	1205	ATGTGTTTCGATGACTTTTACCACTGCAATGCCCAAAAAAGAACGAGAGAGAAAACTCTC	1264
Db	1212	ATGTGTTTCGATGACTTTTACCACTGCAATGCCCAAAAAAGAACGAGAGAGAAAACTCTC	1271
Qy	1265	GCTCGATTTGCCGATGTTACGCAATGAGGACGAGAAATATTCAAGGAAATACACAGAGCTTC	1324
Db	1272	GCTCGATTTGCCGATGTTACGCAATGAGGACGAGAAATATTCAAGGAAATACACAGAGCTTC	1331
Qy	1325	AGGAAAAGTCAATCAATGAGGCAACTTCCAAAAGAGAGCCCTGCACCACTTCACCAACGTC	1384
Db	1332	AGGAAAAGTCAATCAATGAGGCAACTTCCAAAAGAGAGCCCTGCACCACTTCACCAACGTC	1391
Qy	1385	GCAGAAATGCATTTCGCGCATCGCCACCTCTCTAGTCCGCAA-----AG	1426
Db	1392	GCAGAAATGCATTTCGCGCATCGCCACCTCTCTAGTCCGCAAAGGTGAATCTCTCTTCCAG	1451
Qy	1427	CCAGATCCGAGGACTTCGAGGCCAGCAGCCCAATTTCCACTTTGAGCAGCAGGCGGCTCACC	1486
Db	1452	CCAGATCCGAGGACTTCGAGGCCAGCAGCCCAATTTCCACTTCGAGCAGCAGGCGGCTCACC	1511
Qy	1487	AAGGAAGTATGGGCTACATGGAGATATGTACATAGGAAATGATAACAGAGAGAACTCTT	1546
Db	1512	AAGGAAGTATGGGCTACCATGGAGATATGTACATAGGAAATGATAACAGAGAGAACTCTT	1571
Qy	1547	ATCAGGGACACTTTCAAACGCGCATGCGCTTTGACGGTCGACCAATACAGGCCTATATT	1606
Db	1572	ATCAGGGACACTTTCAAACGCGCATGCGCTTTGACGGTCGACCAATACAGGCCTATATT	1631
Qy	1607	ACGTATACGCCCAGATATGCTTACAACTCGCACGACCAAGACGGAATTTATCGCTTTTC	1666
Db	1632	ACGTATACGCCCAGATATGCTTACAACTCGCACGACCAAGACGGAATTTATCGCTTTTC	1691
Qy	1667	AAGGAGACATCTCATCTCTGAGTGTGAAACAGGTGCCCCCAACATGCGCACATAAGG	1726
Db	1692	AAGGAGACATCTCCATCTCTGAGTGTGAAACAGGTGCCCCCAACATGCGCACATAAGG	1751
Qy	1727	TGCACACTGCCACACGAGTGGTCTGATCCACCTGGAAACGAAACGAGAGGATCCATCTGA	1786
Db	1752	TGCACACTGCCACACGAGTGGTCTGATCCACCTGGAAACGAAACGAGAGGATCCATCTGA	1811
Qy	1787	AGGACATTTCACACGATCGCAATGCACTTCTGCGGGAGGGGAAAACACCGAAGCTACTTTG	1846
Db	1812	AGGACATTTCACACGATCGCAATGCACTTCTGCGGGAGGGGAAAACACCGAAGCTACTTTG	1871
Qy	1847	GCATCTTCAAGGTGTAATTTGGAGAGATTTATCCCGGTGAGAGATGGAATACAGATTTA	1906
Db	1872	GCATCTTCAAGGTGTAATTTGGAGAGATTTATCCCGGTGAGAGATGGAATACAGATTTA	1931
Qy	1907	AGCTTTTGTCCCGCGACTCTCGTGAATGCGATTCATCGCCAGCGTGAATCCATTTAGTT	1966
Db	1932	AGCTTTTGTCCCGCGACTCTCGTGAATGCGATTCATCGCCAGCGTGAATCCATTTAGTT	1991
Qy	1967	CGTAGTACCTAGTCTTAGTCACCTCCAAACCTTAATCTCAATCGGAAATCGTCATCTGCAT	2026
Db	1992	CGTAGTACCTAGTCTTAGTCACCTCCAAACCTTAATCTCAATCGGAAATCGTCATCTGCAT	2051
Qy	2027	TAGTCAGAGACGAGGAGAAATCATATTTATTTTGTATATATCTGCTTCGACTCTAAAAAG	2086
Db	2052	TAGTCAGAGACGAGGAGAAATCATATTTATTTTGTATATATCTGCTTCGACTCTAAAAAG	2111



```
Db 1261 AGCAGCGCGCGTCAACAAGGAAGTATGGCTACCATGGAGATATGTACATAGGAAT 1320
Qy 1528 GATACGAGAGAAACTCTTATCAGGAGACATTTCAACCGCGATGGCGTCTTGACGGTG 1587
Db 1321 GATAACGAGAGAAACTCTTATCAGGAGACATTTCAACCGCGATGGCGTCTTGACGGTG 1380
Qy 1588 ACCAATACAGGCTATATATACGTATACGCCAGATATGCTACAACTCGCAGCAGCAG 1647
Db 1381 ACCAATACAGGCTATATATACGTATACGCCAGATATGCTACAACTCGCAGCAGCAG 1440
Qy 1648 AACGATTATATGCTTTCAAGGAGACACTCCATTCCTGCGAGTGTGAAACACGGTGCCC 1707
Db 1441 AACGATTATATGCTTTCAAGGAGACACTCCATTCCTGCGAGTGTGAAACACGGTGCCC 1500
Qy 1708 ACCAATACGACATTAAGGTGCACACTGCGACAGTGGTCTGATCCACTGGAAGA 1767
Db 1501 ACCAATACGACATTAAGGTGCACACTGCGACAGTGGTCTGATCCACTGGAAGA 1560
Qy 1768 AACGAGAGGATCCATCTGAAGGACATTCACAACTGCAATGCAATGCTGCGGAGGGA 1827
Db 1561 AACGAGAGGATCCATCTGAAGGACATTCACAACTGCAATGCAATGCTGCGGAGGGA 1620
Qy 1828 AACAAACGAGAGTACTTTGGGCAATCTCAAGGTGTAA 1863
Db 1621 AACAAACGAGAGTACTTTGGCATCTTCAAGGTGTAA 1656

RESULT 5
AY115551
LOCUS Drosophila melanogaster DARTH (darth) mRNA linear INV 01-AUG-2003
DEFINITION Drosophila melanogaster DARTH (darth) mRNA, complete cds.
ACCESSION AY115551
VERSION AY115551.1 GI:31321973
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1248)
Kaupilla,S., Maaty,W.S., Chen,P., Tomar,R.S., Eby,M.T., Chapiro,J.,
Chew,S., Rathore,N., Zachariah,S., Sinha,S.K., Abrams,J.M. and
Chaudhary,P.M.
Eiger and its receptor, Wengen, comprise a TNF-like system in
Drosophila
Oncogene 22 (31), 4860-4867 (2003)
22775938
MEDLINE 12894227
PUBMED
REFERENCE 2 (bases 1 to 1248)
Chaudhary,P.M.
Direct Submission
TITLE
Submitted (28-MAY-2002) Internal Medicine, UT Southwestern Medical
Center, 5323 Harry Hines Blvd., Dallas, TX 75390, USA
JOURNAL
Location/Qualifiers
FEATURES
1..1248
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/maps="2R"
1..1248
/gene="darth"
1..1248
/gene="darth"
/note="Drosophila aspartate rich TNF homolog; tumor
necrosis factor ligand family member"
/codon_start=1
/product="DARTH"
/protein_id="AAM66763.1"
/db_xref="GI:31321974"
/translation="MTAETLKPFTTPTTSANDGFFPAKATSTATQRRTRQLPIPLVLF
IGLGLVVAIALTIWQTRVSHLDKELSLKRVVDNLQORLGINLYLDFDFQKEYEN
```

```
ALIDPKVKVGLTDEDDDDGDLSDIADDDDDVSYSVDVVDGADYEDYDTMLNKLN
NAHTGTTPTSETTAEGEGETSDASSAGNDNDVFDFTSYNAHKKKQKRSKSIADVRN
EEQNIQGNHTELOKSSNEATSKESPAPLHRRMRHSHRHLLVRKGESLLSARSEDS
RPAAHFHLSSRRHQSGMGVGHDMYIGNDAERNYSQGHFOTRDGVLTVNTGLYVYA
QICYNRSHDQNGFIVFOGDTFPFLQCLNTVPTNMPHKVHTCHTSLGHLERNERHLKD
IHNDRNAVLRGNNRNSYFGIPKV"
1..108
/misc_feature
/gene="darth"
/region="intracellular domain"
109..177
/misc_feature
/gene="darth"
/note="Region: transmembrane domain"
178..1245
/misc_feature
/gene="darth"
/note="Region: extracellular domain"
433..438
/misc_feature
/gene="darth"
/note="cleavage site"

ORIGIN
Query Match 56.0%; Score 1202; DB 3; Length 1248;
Best Local Similarity 98.6%; Pred. No. 1.6e-263;
Matches 1230; Conservative 0; Mismatches 0; Indels 18; Gaps 1;
Qy 634 ATGACTGCGGAGACCCCTCAAGCCGTTTATACGCCAACGAGTGCACGATGATGTTT 693
Db 1 ATGACTGCGGAGACCCCTCAAGCCGTTTATACGCCAACGAGTGCACGATGATGTTT 60
Qy 694 CCGGCCAAAGCGACGACGCGGCGACGCCCGCAGGAGCGCACCCGCGAGCTATCCCTG 753
Db 61 CCGGCCAAAGCGACGACGCGGCGACGCCCGCAGGAGCGCACCCGCGAGCTATCCCTG 120
Qy 754 GTTTTGGGGTTCATCGGTCTGGGGCTGGTCTGCCATTCTCGCACTAACGATCTGGCAG 813
Db 121 GTTTTGGGGTTCATCGGTCTGGGGCTGGTCTGCCATTCTCGCACTAACGATCTGGCAG 180
Qy 814 ACAACGGGTGTATCGCATCTGGACAAGGAGCTGAAGAGCCTGAAGCGAGTCTGCGATAAT 873
Db 181 ACAACGGGTGTATCGCATCTGGACAAGGAGCTGAAGAGCCTGAAGCGAGTCTGCGATAAT 240
Qy 874 CTCAGCAGCGTTTGGGCATTAACATATCTGACAGTTCGACGAGTTCCTCAAAAGGAGTAC 933
Db 241 CTCAGCAGCGTTTGGGCATTAACATATCTGACAGTTCGACGAGTTCCTCAAAAGGAGTAC 300
Qy 934 GAGAAATGCCCTCATCGATCTCAAAAAGGTGATGSCCTCAGGATGAGGAGGAGCAGC 993
Db 301 GAGAAATGCCCTCATCGATCTCAAAAAGGTGATGSCCTCAGGATGAGGAGGAGCAGC 360
Qy 994 GACGATGGCGATGGTCTGGATTCCATTGCGGACGACGAGGAGCAGCGTTAGCTATAGC 1053
Db 361 GACGATGGCGATGGTCTGGATTCCATTGCGGACGACGAGGAGCAGCGTTAGCTATAGC 420
Qy 1054 TCTGTGATGATGTTGGCGCAGACTACGAGAGCTACACCGATATGTTAAATTAACCTCAAC 1113
Db 421 TCTGTGATGATGTTGGCGCAGACTACGAGAGCTACACCGATATGTTAAATTAACCTCAAC 480
Qy 1114 AATCCATATCCGGCACCGACCCACATCTGAGACCACTGCTGAGGGCGGGGAGAGACG 1173
Db 481 AATCCATATCCGGCACCGACCCACATCTGAGACCACTGCTGAGGGCGGGGAGAGACG 540
Qy 1174 GACAGTGCATCCTCAGGCTCAAAATGATGACAAATGTGTTTCGATGACTTTTACCAGCTACAAT 1233
Db 541 GACAGTGCATCCTCAGGCTCAAAATGATGACAAATGTGTTTCGATGACTTTTACCAGCTACAAT 600
Qy 1234 GCCCAAAAAAGAGCGAGGAGAGAAAATCTTCGCTCGATTGCCGATGTACGCAATGAGGAG 1293
Db 601 GCCCAAAAAAGAGCGAGGAGAGAAAATCTTCGCTCGATTGCCGATGTACGCAATGAGGAG 660
Qy 1294 CAGATATTTCAAGGAATTCACAGAGCTTCAGGAAAGTCAATCCATGAGGCACTTCC 1353
Db 661 CAGATATTTCAAGGAATTCACAGAGCTTCAGGAAAGTCAATCCATGAGGCACTTCC 720
Qy 1354 AAGAGAGCCCTGCACCACTTTCACCAACCGTCGAGAATGCAATTCGCCCATCGCCACCTC 1413
```

721 AAAGAGAGCCCTGCACCACTTCACCAAGCTCGCAGATGCATTCGCCCATGCCACCTC 780  
1414 CTAGTCGGCAA-----AGCAGATCGAGACTCGAGGCGAGCAGCC 1455  
781 CTAGTCGGCAAAGGTGAATCTCTTTTTCAGGCAGATCGAGGACTCGAGGCGAGCAGCC 840  
1456 CATTTCCACTTGAGCAGCAGCGCGCTCAACAGGAAGTATGGCTACCATGGAGATATG 1515  
841 CATTTCCACTTGAGCAGCAGCGCGCTCAACAGGAAGTATGGCTACCATGGAGATATG 900  
1516 TACATAGGAATGATTAACAGAGAAACTCTTATCAGGGACACTTTCAAACGCGCGATGGC 1575  
901 TACATAGGAATGATTAACAGAGAAACTCTTATCAGGGACACTTTCAAACGCGCGATGGC 960  
1576 GTCTTCAGCGTGACCAATACAGCGCTATATTAGTATACGCCACGATATGCTAACAAAC 1635  
961 GTCTTCAGCGTGACCAATACAGCGCTATATTAGTATACGCCACGATATGCTAACAAAC 1020  
1636 TCGCAGCAGCAGAAAGGATTTATCGTCTTTCAAGGAGACACTTCCATTCCTGCAGTCTTG 1695  
1021 TCGCAGCAGCAGAAAGGATTTATCGTCTTTCAAGGAGACACTTCCATTCCTGCAGTCTTG 1080  
1696 ACACGGTGTCCACCAACATGCCACATAGGTGCACTCGCCACAGAGTGTCTGATC 1755  
1081 AACACGGTGTCCACCAACATGCCACATAGGTGCACTCGCCACAGAGTGTCTGATC 1140  
1756 CACCTTGGAAACGAAACGAGAGGATTCATCTGAGGACATTCACAGATCGCAATGCAGTT 1815  
1141 CACCTTGGAAACGAGAGGATTCATCTGAGGACATTCACAGATCGCAATGCAGTT 1200  
1816 CTGCGGAGGGAAACAAACCGAAGTACTTTGGCATCTTCAAGGTGTAA 1863  
1201 CTGCGGAGGGAAACAAACCGAAGTACTTTGGCATCTTCAAGGTGTAA 1248

RESULT 6  
AY119233  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AY119233  
Drosophila melanogaster 1221 bp mRNA linear INV 16-JUN-2002  
AY119233  
AY119233.1 GI:21430829  
FLI CDNA.  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1221)  
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,  
Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E.,  
George, R., Gonzalez, M., Guarin, H., Kronmiller, B., Li, P., Liao, G.,  
Miranda, A., Mungall, C.J., Nunco, J., Pacleb, J., Paragas, V., Park, S.,  
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Rubin, G.M.  
and Celniker, S.  
Direct Submission  
Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,  
Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, USA  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720  
This clone was sequenced as part of a high-throughput process to  
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
Science 2000). The sequence has been subjected to integrity checks  
for sequence accuracy, presence of a polyA tail and contiguity  
within 100 kb in the genome. Thus we believe the sequence to  
reflect accurately this particular cDNA clone. However, there are  
artifacts associated with the generation of cDNA clones that may  
have not been detected in our initial analyses such as internal  
priming, priming from contaminating genomic DNA, retained introns  
due to reverse transcription of unspliced precursor RNAs, and

reverse transcriptase errors that result in single base changes.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our web site  
(http://fruitfly.berkeley.edu) or send email to  
cdna@fruitfly.berkeley.edu.

FEATURES  
source  
1..1221  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
1..1221  
/gene="CG12919"  
/db\_xref="FLYBASE:FBgn0033483"  
170..955  
/genes="CG12919"  
/notes="Longest ORF"  
/codon\_start=1  
/product="SD18286p"  
/protein\_id="AAM51093.1"  
/db\_xref="GI:21430830"  
/db\_xref="FLYBASE:FBgn0033483"  
/translation="MLNKLNNARHTGTPITSETTAEGEGTDSASSASNDNVDFDPTS  
YNAHKKQERKSRSIADVNERONIQNHTELQEKSSNEATSKSPALPHRRRMRHSHR  
HRLHLVRKESLILSARSDSRPAAPHLSRRRHQSGMGTGMDTNGDNRNRSYQGH  
FQTRDGLVNTNGLYYVYQICYNNSHDQNGFIYFGDTPFLQCLNTVPTNMPHKVH  
TCHTSLIHLERNERLHLDKIHDRNVAFLREGRNRSYFGIFKV"

ORIGIN  
Query Match 53.5%; Score 1149.4; DB 3; Length 1221;  
Best Local Similarity 97.9%; Pred. No. 1.7e-251;  
Matches 1192; Conservative 0; Mismatches 6; Indels 19; Gaps 2;  
QY 931 TAGGAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCTCACGGATGAGGAGAC 990  
DB 6 TAGGAATGCCCTCATCGACTATCC-CAAAAAGGTGGATGGCTCACGGATGAGGAGAC 64  
QY 991 GACGACATGGCGATGGTCTGGATTCATTCGGGACGAGGACGACGATTTAGCTAT 1050  
DB 65 GACGACATGGCGATGGTCTGGATTCATTCGGGACGAGGAGGACGACGATTTAGCTAT 124  
QY 1051 AGCTCTGGATGATGTTGGCGCAGACTACCGGACTACCGGATATGTTAAATAAATC 1110  
DB 125 AGCTCTGGATGATGTTGGCGCAGACTACCGGACTACCGGATATGTTAAATAAATC 184  
QY 1111 AACATGACATACCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 1170  
DB 185 AACATGACATACCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 244  
QY 1171 ACGGACATGTCATCCTCAGCCTCAATGATGACATGTTGTCGATGCTTTACGACGCTAC 1230  
DB 245 ACGGACATGTCATCCTCAGCCTCAATGATGACATGTTGTCGATGCTTTACGACGCTAC 304  
QY 1231 AATGCCCAAAAAAGACGAGGAGAAAATCTCGCTCGATTGCGGATGATGACGAATGAG 1290  
DB 305 AATGCCCAAAAAAGACGAGGAGAAAATCTCGCTCGATTGCGGATGATGACGAATGAG 364  
QY 1291 GAGCAGATATTCAGGGAATACACAGAGCTTCAGGAAAGTATCAATGAGGCAACT 1350  
DB 365 GAGCAGATATTCAGGGAATACACAGAGCTTCAGGAAAGTATCAATGAGGCAACT 424  
QY 1351 TCCAAAGAGAGCCCTGCACCACTTCCACCGTCGAGAAATGTCATTCGCCCATTCGCCAC 1410  
DB 425 TCCAAAGAGAGCCCTGCACCACTTCCACCGTCGAGAAATGTCATTCGCCCATTCGCCAC 484  
QY 1411 CTCCTAGTCCGCAA-----AGCAGATCCGAGGACTCGAGGCCAGCA 1452  
DB 485 CTCCTAGTCCGCAAAGGTGAATCTCTTTTTCAGCCAGATCCGAGGACTCGAGGCCAGCA 544  
QY 1453 GCCCATTTCCACTTTGAGCAGCGCGCTCACCAAGGAAGTATGGCTACCAATGGAGAT 1512  
DB 545 GCCCATTTCCACTTTGAGCAGCGCGCTCACCAAGGAAGTATGGCTACCAATGGAGAT 604  
QY 1513 ATGTACATAGGAATGATTAACGAGAGAAATCTTTATCAGGAGAACTTTTCAACGCGCAT 1572

Db 605 ATGTACATAGGAATGATAACGAGAGAAATCTTATCAGGGACACTTTCAAACGCGCAT 664  
QY GCGCTCTTGACGGTGACCAATACAGGCTATATATTACGTATAGCCGACAGATATGCTACAAAC 1632  
Db 665 GCGCTCTTGACGGTGACCAATACAGGCTATATATTACGTATAGCCGACAGATATGCTACAAAC 724  
QY 1633 AACTGCGACGACGAAACGGATTTATCGTCTTTTCAAGGAGACACTCCCAATCTCTGAGTGC 1692  
Db 725 AACTGCGACGACGAAACGGATTTATCGTCTTTTCAAGGAGACACTCCCAATCTCTGAGTGC 784  
QY 1693 TTGAACACGGTGCCCAACCAATGSCACATAGGTGACACCTGSCACACAGATGCTGTG 1752  
Db 785 TTGAACACGGTGCCCAACCAATGSCACATAGGTGACACCTGSCACACAGATGCTGTG 844  
QY 1753 ATCCACCTGGAAACGAAACGAGAGGATCCATCTGAAGGACATTTCAACACGATCGCAATGCA 1812  
Db 845 ATCCACCTGGAAACGAAACGAGAGGATCCATCTGAAGGACATTTCAACACGATCGCAATGCA 904  
QY 1813 GTTCTGCGGAGGAAACCAACGAAAGTACTTTTGGCATCTTTCAAGGTGTAATTTGGAGAG 1872  
Db 905 GTTCTGCGGAGGAAACCAACGAAAGTACTTTTGGCATCTTTCAAGGTGTAATTTGGAGAG 964  
QY 1873 ATTATCCCGGTGAGAAATGGAATACCAAGTTTAAGCTTTTGTCCCGGACTGCTGTG 1932  
Db 965 ATTATCCCGGTGAGAAATGGAATACCAAGTTTAAGCTTTTGTCCCGGACTGCTGTG 1024  
QY 1933 AATGCGATTCTCGCACGCGTGAATCCATTAGTTCTGTAGTACCTAGTCTTAGTCACTCCA 1992  
Db 1025 AATGCGATTCTCGCACGCGTGAATCCATTAGTTCTGTAGTACCTAGTCTTAGTCACTCAT 1084  
QY 1993 AACCTAATCTCAATCGGAATCGTGATCTACTGATCTAGTCAGTGTGAGGAAATCATATA 2052  
Db 1085 AACCTAATCTCAATCGGAATCGTGATCTACTGATCTAGTCAGTGTGAGGAAATCATATA 1144  
QY 2053 TTTATTTTGTATATCTCGTCTGACTCTAAAGTGAATATAATATATAGTACTATTAA 2112  
Db 1145 TTTATTTTGTATATCTCGTCTGACTCTAAAGTGAATATAATATATAGTACTATTAA 1204  
QY 2113 AAAAAAAAAAAAAA 2129  
Db 1205 AAAAAAAAAAAAAA 1221

RESULT 7  
AC012854  
LOCUS AC012854 55359 bp DNA linear HTG 03-NOV-1999  
DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
AC012854  
VERSION AC012854.1 GI:6223467  
KEYWORDS HTG; HTGS\_PHASE2.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 55359)  
Adams, M. and Venter, J.C.  
Direct Submission  
Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
This sequence was identified as CDM:10210834 by the submitter.  
For further information on this sequence you may e-mail to  
fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
Location/Qualifiers  
1..55359  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"

ORIGIN  
Query Match 37.2%; Score 798.8; DB 2; Length 55359;  
Best Local Similarity 99.8%; Pred. No. 5.9e-171;  
Matches 800; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 5 CGAGCGAAGCGAGCGTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64  
Db 28568 CGAGCGAAGCGAGCGTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 28627  
QY 65 CGCGTGGTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTTCAGTTTAACTTACCAAG 124  
Db 28628 CGCGTGGTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTTCAGTTTAACTTACCAAG 28687  
QY 125 ATTGAATTTTCCCATCGGCAAAATTTACTTAAATAATACATAGTGCAACTCGTCCACTGTGT 184  
Db 28688 ATTGAATTTTCCCATCGGCAAAATTTACTTAAATAATACATAGTGCAACTCGTCCACTGTGT 28747  
QY 185 GTTGTGTTTTTTTTTTTTTTTTTTTTTGGTTTTTTCGCTGTGCTTTTATCGCAAAACGAACATGAT 244  
Db 28748 GTTGTGTTTTTTTTTTTTTTTTTTTTTGGTTTTTTCGCTGTGCTTTTATCGCAAAACGAACATGAT 28807  
QY 245 AAAACTAGAAATATCTTGAGAAACTTTGTTTTCGCGCTTTTCTTTTGTCTAATTCGCCATC 304  
Db 28808 AAAACTAGAAATATCTTGAGAAACTTTGTTTTCGCGCTTTTCTTTTGTCTAATTCGCCATC 28867  
QY 305 GCGGAAGAGAAACAAACAGCAGTAGACAAACAAAGTGTGTAATAACAATCTGAAAGGGCA 364  
Db 28868 GCGGAAGAGAAACAAACAGCAGTAGACAAACAAAGTGTGTAATAACAATCTGAAAGGGCA 28927  
QY 365 CCATCAGACGCCGAGGGGTTTTATCTATATAGATGTCCGAGCTTATCATCTCATCTGCTGTC 424  
Db 28928 CCATCAGACGCCGAGGGGTTTTATCTATATAGATGTCCGAGCTTATCATCTCATCTGCTGTC 28987  
QY 425 TGTGAGGTTGTTCTGTGTGCTGCTCGTAGTATCTTAAATACATAGAGTGTGTTTATATAAA 484  
Db 28988 TGTGAGGTTGTTCTGTGTGCTGCTCGTAGTATCTTAAATACATAGAGTGTGTTTATATAAA 29047  
QY 485 GTGCGACAAAGCTCGATTGGAAACAGCTGTGAGTGCCCTTGAGTGGTGGGCAAGATCG 544  
Db 29048 GTGCGACAAAGCTCGATTGGAAACAGCTGTGAGTGCCCTTGAGTGGTGGGCAAGATCG 29107  
QY 545 TCATCATCATCATCGTCTGCTCATTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGGT 604  
Db 29108 TCATCATCATCATCGTCTGCTCATTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGGT 29167  
QY 605 TGCTCTAAGATCCCGAGTGTTCATCAATATAGTCTGCGAGACCCCTCAAGCCGTTTATAA 664  
Db 29168 TGCTCTAAGATCCCGAGTGTTCATCAATATAGTCTGCGAGACCCCTCAAGCCGTTTATAA 29227  
QY 665 CGCCACGAGTGCCCAACGATGATGTTTTCGGGCAAGCGACGACGACGCGCCGCC 724  
Db 29228 CGCCACGAGTGCCCAACGATGATGTTTTCGGGCAAGCGACGACGACGCGCCGCC 29287  
QY 725 AGCGACGACCCCGCAGCTGATCCCGTGGTTTTTGGGGTTTATCGGTTCTGGGGCTGTGTCG 784  
Db 29288 AGCGACGACCCCGCAGCTGATCCCGTGGTTTTTGGGGTTTATCGGTTCTGGGGCTGTGTCG 29347  
QY 785 TTGCCATTCTCGCACTAACGAT 806  
Db 29348 TTGCCATTCTCGCACTAACGAT 29369  
RESULT 8  
AC005974/c  
LOCUS AC005974 158983 bp DNA linear INV 19-NOV-1998  
DEFINITION Drosophila melanogaster, chromosome 2R, region 46E1-46F6, P1 clones  
DS05033, DS01913, and DS05181, complete sequence.  
AC005974 AC005433 AC005436 AC005469  
AC005974.1 GI:3893034  
HTG.  
KEYWORDS Drosophila melanogaster (fruit fly)  
SOURCE Drosophila melanogaster

REFERENCE AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 158983)
REFERENCE AUTHORS	Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E., Svirkas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.
TITLE JOURNAL REFERENCE AUTHORS	Sequencing of Drosophila chromosome 2R, region 46E1-46F6 Unpublished (1998) 2 (bases 1 to 158983) Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E., Svirkas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.
TITLE JOURNAL	Direct Submission
COMMENT	Submitted (19-NOV-1998) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US Sequence submitted by: Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720
FEATURES source	For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site ( <a href="http://www.fruitfly.org/sequence/">http://www.fruitfly.org/sequence/</a> ) or send email to <a href="mailto:bdgp@fruitfly.berkeley.edu">bdgp@fruitfly.berkeley.edu</a> . PI library locations: 53-41, 20-89, 54-93. Location/Qualifiers 1..158983 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /strain="y2; cn bw sp" db_xref="Caxon:7227" /chromosome="2R" /map="46E1-46F6" /clone="P1s DS05033 (D347), DS01913 (D350), and DS05181 (D384)" /clone_lib="P1 library, partial Sau3A in pAd10acBII" /note="These P1s were completed as a project. DS05181 (D384) is a partial bridge with its distal neighbor extending from bp 1 to P1 end at bp 57,624. DS01913 (D350) extends from P1 end at 7,924 to P1 end at 92,484. DS05033 (D347) extends from P1 end at bp 80,881 to P1 end at bp 158,983. Between bps 53,139 and 53,140, Tn10 was excised from DS05181 along with an associated 9 bp duplication."
ORIGIN	
Query Match	37.2%; Score 798.8; DB 3; Length 158983;
Best Local Similarity	99.8%; Pred. No. 7.9e-171;
Matches 800; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	5 CGAGCGCAACGGAGCTTTAAAGTGAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64
Db	134656 CGAGCGCAACGGAGCTTTAAAGTGAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 134597
QY	65 CGCGTGGCGTGCATGATAGCGAAGAAAAGCTATCCGTTTCAGTTAACTACTTACCAAG 124
Db	134596 CGCGTGGCGTGCATGATAGCGAAGAAAAGCTATCCGTTTCAGTTAACTACTTACCAAG 134537
QY	125 ATTGAATTTCCGATCGGGCAATTTACTTAAATAATACATAGTGCAACTCGTCCACTGTGT 184
Db	134536 ATTGAATTTCCGATCGGGCAATTTACTTAAATAATACATAGTGCAACTCGTCCACTGTGT 134477
QY	185 GTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTGGCTGTGCTTTATCGCAAAACAAGACTGAT 244
Db	134476 GTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTGGCTGTGCTTTATCGCAAAACAAGACTGAT 134417
QY	245 AAAACTAGAAAATATCTTGAGAACTTGTGTTTTTTCGCGCTTTTCTTTTCTGCTAAATTCGCGATC 304
Db	134416 AAAACTAGAAAATATCTTGAGAACTTGTGTTTTTTCGCGCTTTTCTTTTCTGCTAAATTCGCGATC 134357
QY	305 GCGGAACAGAAAAACAAGCAGTAGACAAAACAGTGTGGTAATACAACTCTGAAAAGGGCA 364
Db	134356 GCGGAACAGAAAAACAAGCAGTAGACAAAACAGTGTGGTAATACAACTCTGAAAAGGGCA 134297
QY	365 CCATCAGCAGCCCGAGGGTTTTATCTATATAGATGTGCGAGCTTATCATCTCATGTGTC 424
Db	134296 CCATCAGCAGCCCGAGGGTTTTATCTATATAGATGTGCGAGCTTATCATCTCATGTGTC 134237
QY	425 TGTGAGGTGTTCTGTGCTGTGTAGTATCTTAAATACATAGAGTGTGTTTCAATAAA 484
Db	134236 TGTGAGGTGTTCTGTGCTGTGTAGTATCTTAAATACATAGAGTGTGTTTCAATAAA 134177
QY	485 GTCCGACAAAGCTCGATTGGAACACGCTGTCGAGTGCCTTTCAGTGGGTGGCGCAAGATCG 544
Db	134176 GTCCGACAAAGCTCGATTGGAACACGCTGTCGAGTGCCTTTCAGTGGGTGGCGCAAGATCG 134117
QY	545 TCATCATCATCATCTGTCGTCATTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCCGGT 604
Db	134116 TCATCATCATCATCTGTCGTCATTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCCGGT 134057
QY	605 TGCTCTAAGATCCCGAGTGTTCATCAATTCATGCTGCGAGACCCCTCAAGCGGTTTATAA 664
Db	134056 TGCTCTAAGATCCCGAGTGTTCATCAATTCATGCTGCGAGACCCCTCAAGCGGTTTATAA 133997
QY	665 CGCCAACAGTGCCTCAACAGATGATGTTTTTCCGGCCAAAGCCAGACGCGCGCCGCC 724
Db	133996 CGCCAACAGTGCCTCAACAGATGATGTTTTTCCGGCCAAAGCCAGACGCGCGCCGCC 133937
QY	725 ACGGAGCACCCCGCAGCTGATCCCGCTGGTTTTGGGGTTCATCGGTCTGGGGTGGTGTG 784
Db	133936 ACGGAGCACCCCGCAGCTGATCCCGCTGGTTTTGGGGTTCATCGGTCTGGGGTGGTGTG 784
QY	785 TTGCCATTCTCGCACTAACGAT 806
Db	133876 TTGCCATTCTCGCACTAACGAT 133855
RESULT 9	
AC007414	172904 bp DNA linear INV 06-SEP-2001
LOCUS	Drosophila melanogaster, chromosome 2R, region 46C-46D, BAC clone
DEFINITION	BACR06701, complete sequence.
AC007414	
ACCESSION	AC007414.6 GI:15451491
VERSION	HTG.
KEYWORDS	Drosophila melanogaster (fruit fly)
SOURCE	Drosophila melanogaster
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 172904)
REFERENCE	Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Ananides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,E.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirkas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

TITLE		Sequencing of Drosophila chromosome 2R, region 46C-46D	
JOURNAL	Unpublished		
	2 (bases 1 to 172904)		
REFERENCE	Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Svirskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.		
	Direct Submission		
JOURNAL	Submitted (28-APR-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA		
	On Sep 6, 2001 this sequence version replaced gi:13324748.		
COMMENT	Sequence submitted by:		
	Berkeley Drosophila Genome Project		
COMMENT	Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720		
	This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site ( <a href="http://www.fruitfly.org/sequence/">http://www.fruitfly.org/sequence/</a> ) or send email to <a href="mailto:bdgp@fruitfly.berkeley.edu">bdgp@fruitfly.berkeley.edu</a> .		
FEATURES	Location/Qualifiers		
	1..172904		
ORIGIN	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
ORIGIN	/strain="y; cn bw sp"		
	/db_xref="taxon:7227"		
ORIGIN	/chromosomes="2R"		
	/map="46C-46D"		
ORIGIN	/clone="BACR06J01 (D601)"		
	/clone_lib="RPCI-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in pBACE3.6)"		
ORIGIN	Query Match 37.2%; Score 798.8; DB 3; Length 172904;		
	Best Local Similarity 99.8%; Pred. No. 8.1e-171; Matches 800; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	5	CGAGCGCAACGACGCTTTAAAGTCGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG	64
	Db	133226	CGAGCGCAACGACGCTTTAAAGTCGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG
Qy	65	CGCGTGCCTGCATGATAGCGAAGAAAAAGTATCCGTTTCAGTTAACTACTTACCAAG	124
	Db	133286	CGCGTGCCTGCATGATAGCGAAGAAAAAGTATCCGTTTCAGTTAACTACTTACCAAG
Qy	125	ATTGAATTCGCCATCGGCAAAATTAATAAATAATACATAAGTGAACCTGCTCCACTGTGT	184
	Db	133346	ATTGAATTCGCCATCGGCAAAATTAATAAATAATACATAAGTGAACCTGCTCCACTGTGT
Qy	185	GTGTGCTGTTTTTTTTTTTTTTTGGTTTTCGCTGTCCTTTATCGCAACCAAGAACTGAT	244
	Db	133406	GTGTGCTGTTTTTTTTTTTTTTTGGTTTTCGCTGTCCTTTATCGCAACCAAGAACTGAT
Qy	245	AAACTAGAAAATATCTTGAGAAACCTGTTTTCGCGCTTTTCTTTTGCTAATTGCCGATC	304
	Db	133466	AAACTAGAAAATATCTTGAGAAACCTGTTTTCGCGCTTTTCTTTTGCTAATTGCCGATC
Qy	305	CGGGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGGTGAATCAATCTGAAAAGGGCA	364
	Db	133526	CGGGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGGTGAATCAATCTGAAAAGGGCA
Qy	365	CCATCAGACGCCGAGGGGTTTATCTATATAGATGTCGCACGCTATCATCTCATGCTGTC	424
	Db	133586	CCATCAGACGCCGAGGGGTTTATCTATATAGATGTCGCACGCTATCATCTCATGCTGTC
Qy	425	TGTGAGGTTGTTCTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTTATATAAA	484
	Submitted (08-NOV-2001) Berkeley Drosophila Genome Project, MS		



64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory, MS 64-121  
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to [bdgp@fruitfly.berkeley.edu](mailto:bdgp@fruitfly.berkeley.edu).  
Location/Qualifiers

## FEATURES

Source

1. .189620  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/strain="y; cn bw sp"  
/db\_xref="taxon:7227"  
/chromosome="2R"  
/map="46D-47A"  
/clone="BACR09J12 (D1345)"  
/clone.lib="RP11-98 (Roswell Park Cancer Institute  
Drosophila melanogaster BAC library, partial EcoRI in  
pBACE3.6)"

## ORIGIN

Query Match 37.2%; Score 798.8; DB 3; Length 189620;  
Best Local Similarity 99.8%; Pred. No. 8.3e-171;  
Matches 800; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 CGAGCGCAACGGACGTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64  
DB 169512 CGACGCAAGCGGACGTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 169453  
  
QY 65 CGCGTGGCTGATGATGACGAGAAAAGAAAGCTATCCGTTTCAGTTAACTTACTTACCAAG 124  
DB 169452 CGCGTGGCTGATGATGACGAGAAAAGAAAGCTATCCGTTTCAGTTAACTTACTTACCAAG 169393  
  
QY 125 ATTGAATTTCCGATCGGGCAATTTACTTAAATAATACATAGTGAACCTCGTCCACTGTGT 184  
DB 169392 ATTGAATTTCCGATCGGGCAATTTACTTAAATAATACATAGTGAACCTCGTCCACTGTGT 169333  
  
QY 185 GTTGTGTTTTTTTTTTTTTTTTTGGTTTTCGCTGTGCTTTTATCGCAAAACAAGACTGAT 244  
DB 169332 GTTGTGTTTTTTTTTTTTTTTTTGGTTTTCGCTGTGCTTTTATCGCAAAACAAGACTGAT 169273  
  
QY 245 AAAAAGTAAATATCTTGAGAACTGTTTTTCGCGTTTCTTTTCGCTAATTCGCGATC 304  
DB 169272 AAAAAGTAAATATCTTGAGAACTGTTTTTCGCGTTTCTTTTCGCTAATTCGCGATC 169213  
  
QY 305 GCGGAAGAGAAAAACAAGCAGTAGACAAAACAAGTGTGTAATACATCTGAAAAGGCA 364  
DB 169212 GCGGAAGAGAAAAACAAGCAGTAGACAAAACAAGTGTGTAATACATCTGAAAAGGCA 169153  
  
QY 365 CCATCAGCAGCCGAGGGTTTATCTATATAGATGTCGAGCTTATCATCTCATGCTGTC 424  
DB 169152 CCATCAGCAGCCGAGGGTTTATCTATATAGATGTCGAGCTTATCATCTCATGCTGTC 169093  
  
QY 425 TGTGAGGTGTTGTCGTGCTCGTGTAGTATCTTAAATACATAGAGTGTTTATATAA 484  
DB 169092 TGTGAGGTGTTGTCGTGCTCGTGTAGTATCTTAAATACATAGAGTGTTTATATAA 169033  
  
QY 485 GTGCGAACAAGCTCGATTGGAACAAGCTGTGAGTGCCCTTTGAGTGGGTGGGCAAGATCG 544  
DB 169032 GTGCGAACAAGCTCGATTGGAACAAGCTGTGAGTGCCCTTTGAGTGGGTGGGCAAGATCG 168973  
  
QY 545 TCATCATCATCGTCTGTCATTATCAAGAGATCAGCATCTCGAGGCCCGCGGT 604  
DB 168972 TCATCATCATCGTCTGTCATTATCAAGAGATCAGCATCTCGAGGCCCGCGGT 168913  
  
QY 605 TGCTCTAAGATCCCGAGTGTTCATCAATTATGACTGCGAGACCCCTCAAGCGTTTATATA 664  
DB 168912 TGCTCTAAGATCCCGAGTGTTCATCAATTATGACTGCGAGACCCCTCAAGCGTTTATATA 168853

QY

665 CGCAACGAGTGCACACGATGATGGTTTTCGGCCAAAGCAGCAGCAGCGCCGCC 724

Db

168852 CGCAACGAGTGCACACGATGATGGTTTTCGGCCAAAGCAGCAGCAGCGCCGCC 168793

QY

725 AGGAGCGACCGCCAGCTGATCCCTGGTTTGGGTTTCATCGTCTGGGCTGGTCG 784

Db

168792 AGGAGCGACCGCCAGCTGATCCCTGGTTTGGGTTTCATCGTCTGGGCTGGTCG 168733

QY

785 TTGCGATTCTCGCACTAACGAT 806

Db

168732 TTGCGATTCTCGCACTAACGAT 168711

## RESULT 11

AE003831

LOCUS

AE003831 275390 bp DNA linear INV 14-FEB-2003  
Drosophila melanogaster chromosome 2R, section 19 of 74 of the  
complete sequence.

ACCESSION

AE003831 AE002787 AE013599

VERSION

AE003831.3 GI:21645483

KEYWORDS

Drosophila melanogaster (fruit fly)

SOURCE

Drosophila melanogaster

ORGANISM

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

## REFERENCE

1 (bases 1 to 275390)  
Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,  
George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,  
Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,  
Brandon,R.C., Rogers,C., Blazer,V.R., Blazer,J.G., Champe,M., Pfeiffer,B.D.,  
Wan,K.H., Doyle,C., Baxter,S.G., Helt,G., Nelson,C.R., Gabor,G.L.,  
Abril,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D.,  
Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M.,  
Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S.,  
Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotter,P.,  
Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A.,  
Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B.,  
Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I.,  
Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S.,  
Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C.,  
Ferreira,S., Fleischmann,W., Fosler,C., Gabriellian,A.E., Garu,N.S.,  
Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,  
Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,  
Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,  
Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,  
Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,  
Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A., Li,J.,  
Li,Z., Liang,Y., Lin,X., Liu,X., Mattel,B., McIntosh,T.C.,  
McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,  
Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,  
Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,  
Nusker,D.R., Pacle,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,  
Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,  
Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,  
Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,  
Stapleton,M., Strong,R., Sun,B., Svirskas,R., Tector,C., Turner,R.,  
Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,  
Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T.,  
Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,  
Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,  
Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,  
Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.  
The genome sequence of Drosophila melanogaster  
Science 287 (5461), 2185-2195 (2000)  
20196006  
10731132  
2 (bases 1 to 275390)  
Celniker,S.E., Adams,M.D., Kronmiller,B., Wan,K.H., Holt,R.A.,  
Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C., Rogers,Y.,  
Banzon,J., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,

Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Friese,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houch,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Maccel,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirskas,R., Tector,C., Tyler,D., Williams,S.M., Zaveri,J.S., Smith,H.O., Venter,J.C. and Rubin,G.M. Sequencing of *Drosophila melanogaster* genome Unpublished  
3 (bases 1 to 275390)  
Miera,S., Crosby,M.A., Matthews,B.B., Bayraktaroglu,L., Campbell,K., Hradscky,P., Huang,Y., Kaminker,J.S., Prochnick,S.E., Smith,C.D., Tupy,J.L., Bergman,C.M., Berman,B.P., Carlson,J.W., Celisner,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Friese,E., de Grey,A.D.N.J., Harris,N.L., Kronmiller,B., Marshall,B., Milburn,G.H., Richter,J., Russo,S., Searle,S.M.J., Smith,E., Shu,S., Smutniak,F., Whitfield,E.J., Ashburner,M., Gelbart,W.M., Rubin,G.M., Mungall,C.J. and Lewis,S.E.  
Annotation of *Drosophila melanogaster* genome Unpublished  
4 (bases 1 to 275390)  
Adams,M.D., Celisner,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J. Direct Submission  
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
5 (bases 1 to 275390)  
FlyBase  
Direct Submission  
Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA  
6 (bases 1 to 275390)  
FlyBase  
Direct Submission  
Submitted (23-JAN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA  
On Jul 1, 2002 this sequence version replaced gi:10727672.  
Location/Qualifiers  
1. .275390  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/chromosome="2R"  
/note="genotype: y[1]; cn[1] bw[1] sp[1]; Rh6[1]"  
197. .2954  
/locus\_tag="CG1671"  
/note="last curated on Mon Feb 11 17:53:43 PST 2002"  
/map="46B7-46B9"  
/db\_xref="FLYBASE:FBgn0033454"  
join(197. .286,346. .2954)  
/locus\_tag="CG1671"  
/product="CG1671-RA"  
/db\_xref="FLYBASE:FBgn0033454"  
join(261. .286,346. .2683)  
/locus\_tag="CG1671"  
/codon\_start=1  
/note="CG1671 gene product"  
/protein\_id="AAF58888.1"  
/db\_xref="GI:7303841"  
/db\_xref="FLYBASE:FBgn0033454"  
/translation="MLAGISEKSYAAEARVTFNYAGDIAWSPNGHHLYCLINGSVAVNQ  
VDTQSLNYSGLSSPSADNKRPAEDLNIDVEDTITCGLSQESHLVTAHRSGLLR  
WLDGTGLVKLWAKHQGPVIRVFSPGRLICTGGADATRLWDYNNNSCLGALKD  
PPGAWLLVHPNVLRKVIYAGGSDNTVYAWNYETKLVMKRGHLSONVTGLSPKNA  
SDCNEIVTSRDKVNIWQEQESKQLKVLPLYDELEGVYVADEQLIIVASGSK  
LQQVDPKSWKIRILSQSDFOISRLHCKSELQALVTTQGNILVYDTELEPTFKHLV  
GYNDEILDMCFMGHDFYLAVATNSKHFPLYDTEHDMNCKLIIVGHSPTVMSLAOSNL  
LISVGKDCSVRLMLQDKCSLEALTQQAANCHTSTTCCVAMTHNGATGFSVSDGS

TKVQLVRSKEDRNSYSFNLRYYAALSHDKVEVNCVAPANNKLIATASQDKTAKWLAE  
SNTLQGLGHTGRVMSRPSPDQIVLTSSQCTLRISWISNFSCTIKRFDQECTILR  
AEFLDVGKRTIISAASDGLLKNWIKNTCQLSDIENDRVYALAVSARSNRPFTYGA  
DSKLIRPGDVTQTRNEALDKROANLQEQTLSHLHAQQLHKAFVLALNLDKPKAS  
FDIINFPVRKDEPQQLVDQNVORVALLOHKAWTNRSRHSQANGMILKYLIGD  
ALDDPKARFYNGNMVEVTFPYQVRFKRVTELNKDLAFLEFIVKCM"  
complement (2844. .13679)  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/note="last curated on Tue Mar 12 18:58:27 PST 2002"  
/map="46B9-46B12"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341,  
6647. .8156,8220. .8346,13587. .13679))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RE"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341,  
5819. .5866,6647. .8156,8220. .8346,13587. .13679))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-R1"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341,  
6647. .8156,8220. .8346,12506. .12588,12661. .12893))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RD"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341,  
5819. .5866,6647. .8156,8220. .8346,12506. .12893))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RA"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341,  
5819. .5866,6647. .8156,8220. .8346,12506. .12893))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RJ"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341,  
6647. .8156,8220. .8346,8412. .8546))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RB"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341,  
5819. .5866,6647. .8156,8220. .8346,8412. .8546))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RK"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341,  
6647. .8156,8220. .8544))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RG"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341,  
5819. .5866,6647. .8156,8220. .8544))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RL"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(3273. .3521,3590. .4619,4684. .5341,  
6647. .8156,8220. .8318))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"

gene  
mRNA  
mRNA  
mRNA  
mRNA  
mRNA  
mRNA  
mRNA  
mRNA  
mRNA  
mRNA  
CDS  
Query Match 37.2%; Score 798.8; DB 3; Length 275390;  
Best Local Similarity 99.8%; Pred. No. 9.3e-171;

[illegible]

AUTHORS	TITLE	JOURNAL MEDLINE PUBMED REFERENCE	AUTHORS TITLE JOURNAL	FEATURES SOURCE
---------	-------	---	-----------------------------	--------------------

Takeuchi, K., Hatanaka, A., Kimura, M., Seki, N., Kimura, I., Yamada, S. and Yamashita, S.  
Aspelin, a Novel Extremely Aspartic Acid-rich Protein in Fish Muscle, Promotes Iron-mediated Demethylation of Trimethylamine-N-oxide  
J. Biol. Chem. 278 (48), 47416-47422 (2003)  
22995865  
131229916  
2 (bases 1 to 2040)  
Takeuchi, K., Yamashita, S. and Yamada, S.  
Direct Submission  
Submitted (14-AUG-2003) Kazuharu Takeuchi, Nippon Suisan Kaisha, Ltd., Central Research Laboratory; Kitanomachi, Hachioji, Tokyo 192-0906, Japan (E-mail: k-takeuchi@nissui.co.jp, Tel:81-426-56-5195, Fax:81-426-56-5188)

source

gene  
CDS

```
/organism="Theragra chalcogramma"  
/mol_type="mRNA"  
/db_xref=taxon:48550"  
/tissue_type=skeletal muscle"  
/country=Japan: Hokkaido, Hakodate"  
1..2040  
/gene=aspolin2"  
451..1494  
/gene=aspolin2"  
/codon_start=1  
/product=aspartic acid-rich protein aspolin2"  
/protein_id=BAC87890.1"  
/db_xref=GI:35210494"  
/translation=MASYKTWLLGLLVALLCFQTPLAGVKAQLD  
DDDDDDAGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD  
DDDDDDGADGGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD  
DDDDDDHDEAAGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD  
DDDDDDHHKGDGDKPCESGSDTSEHCDCCKMNCFCVCPCMCTLTCT  
YKVTVADFVDPEKY
```

## ORIGIN

Query Match 3.1%; Score 67.2; DB 5; Length 2040;  
Best Local Similarity 50.6%; Pred. No. 0.00031;  
Matches 162; Conservative 0; Mismatches 158; Indels 0

[illegible]

RESULT 13  
AB117517  
LOCUS  
DEFINITION

AB117517 2235 bp mRNA linear VRT 27-NOV-2003  
Theragra chalcogramma aspolin1 mRNA for aspartic acid-rich protein  
aspolin1, complete cds.



JOURNAL Patent: WO 0300899-A 5263 03-JAN-2003;  
Syngenta Participations AG (CH)  
FEATURES Location/Qualifiers  
source 1..2000  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"

Qy 1610 TATA 1613  
Db 905 TCWW 908

Search completed: September 15, 2004, 00:00:18  
Job time : 13370 secs

## ORIGIN

Query Match	3.1%	Score 66.4	DB 6	Length 2000
Best Local Similarity	10.0%	Pred. No. 0.00046		
Matches	90	Conservative 415	Mismatched 391	Indels 8
Qy	710	GCACGGCAGCCCGCAGCGAGCAGCACCCGCCAGCTGATCCCCTGGTTTGGGGTTCATCG	769	
Db	13	SSWRGRRMYWAGMMSCHARMGSSRMSRWKGSYKRYKSSCGKCKMTTTRRSKWYSASSA	72	
Qy	770	GTCTGGGCTGCTGCTTGCCTTCTCGCACTAACGATCTGCAGACACCGGCTGTATCGC	829	
Db	73	SGRTGSKWSSGSYSGKMKRYKRSKRWRGRGRGRMRMRWRMRRCARSRMRAGGS	132	
Qy	830	ATCTGGACAAGAGCTGAAGAGCTGAAGCGAGCTGTCGATAATCTCCAGCAGCGTTTGG	889	
Db	133	GRMMGGKSRMSYMYWCYARGCGSKRKS KGSWGKTCRRGARGSGWSGAKYKSGMS	192	
Qy	890	GCATAAACATCTCGACGAGTTCGACGAGTTCCAAAAGGAGTACGAGAAATCCCTCATCG	949	
Db	193	KRMWSSCGRSGGRSAYSIYGTSRKGYTKMYTYSASRCNRA YMTTYSYMACSSYT	252	
Qy	950	ACTATCCAAAAGGTGATGGCTTCACGGATGAGGAGGACGACGATGGCGCATGGTC	1009	
Db	253	WCRSKRRSMWKMWRKMRWSRSYGWYSWYKWMCTAYKKSYSRWCMYVGGWRGATR	312	
Qy	1010	TGGATTCCATTGGCGACGACGAGGACGACGAGTGTAGCTATGCTCTGTGGATGATGTTG	1069	
Db	313	YWGGRYSRMAWMTYKMYTYGKMGKRWGAGRMWMSRCWRSKACYYWRMRWRMTTR	372	
Qy	1070	GCGCAGACTACGAGGACTACACCCATATGTTAAATAAAGTCAACAAATGCACATACGGCA	1129	
Db	373	RRWAKKSRPSTRKKRWKRWKRYKMRGYSRMRSCRRARWKRKSGRGAWKGGRGCM	432	
Qy	1130	CCAGCCCAACTCTGAGACCACTGCTGAGGCGGAGGCGGAGACGACAGTGCATCTCTCAG	1189	
Db	433	TCRMKSYGMWRKWSKRMASKYKWSRWYRWRKKCSRTTWMGKTRGMMGTWGRCRYKK	492	
Qy	1190	CCTCAAAATGATGACAAATGTTGTCATGACATTTACCAGCTCAATGCCCCACAAAAGAACG	1249	
Db	493	RSGMKRCRRRRWGRMRWKRKYMSAYTMYRCARKK---YSYSAARKARCWYRGKG	548	
Qy	1250	AGGAGAAAATCTCGCTCGATTCGCCATGTAGCAATGAGGACGAGCAATATTCAGGAA	1309	
Db	549	YYWAGMWRKRYKMYMYKMMWYKRYKSKYSWYKMSYASCKWSARKAGKCKRKSMS	608	
Qy	1310	ATCACACAGAGCTTCAGAAAAGTCATCCAAATGAGGCAACTCCAAAGAGAGCGCTCGAC	1369	
Db	609	AWSKMSRPSRKRCASKRSKSAKRYAMMGWMTSGSRMSRWKSYCYWRKWSGMSKSTCWM	668	
Qy	1370	CACCTTACCACCGTCGCGAGATGCAATCCCGCCATCGCCACCTCTAGTCCGCAAGACCA	1429	
Db	669	YMSKYTYAKYGSYWRYYRAW-CMYMWRYYYYRYSYMTYMAWYTSSTRMAMTKMYSG	727	
Qy	1430	GATCCGAGGACTCGAGCCAGCAGCCATTTCCATTTGACGACGCGCGCTCACAAG	1489	
Db	728	RYWTSWYKYCKSKWYSWYYSWNAKWTWWRRYATRMWMMWYRYSKMYTWTCTWVG	787	
Qy	1490	GAAGTATGGGCTACCATGGAGATATGTGACATAGGAAATGATACGAGAGAACTTTATC	1549	
Db	788	YMWYWRTYMKRMYMYKCTKTYWYSATVYTGTAWA--WMAKTKMRMGMTGAKTRGR	844	
Qy	1550	AGGGACACITTTCAAACGCGCGATGGCGCTTGTAGCGGTGACCAATACAGGCCTATATTAG	1609	
Db	845	ARKARYWKAATWCACTKMWTKGKGAKWATWMAKAWRYKYWSWMAWYTYYYKTRTRYK	904	



PA	(GUAN/) GUAN B.	QY	541	ATCGTCATCATCATCGTCGTCTATCAACAGATCAGCATCAGCATCTCGAGGCC	600
PA	(BOWE/) BOWEN M A.	Db	541	ATCGTCATCATCATCGTCGTCTATCAACAGATCAGCATCAGCATCTCGAGGCC	600
XX					
XX	Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;	QY	601	CGGTTGCTCTAAGATCCCAAGTGTTCATCAATATGACTGCCGAGACCCCTCAAGCCGTTT	660
DR	WPI; 2002-195121/25.	Db	601	CGGTTGCTCTAAGATCCCAAGTGTTCATCAATATGACTGCCGAGACCCCTCAAGCCGTTT	660
XX	P-PSDB; AAU77718.				
PT	New Drosophila tumor necrosis factor molecule, useful in controlling	QY	661	ATAACGCCAAGAGTGCCAAAGATGATGTTTTCGGGCCAAAGCGACGACGCGGACC	720
PT	agriculturally important pests, e.g. comprises modifying the growth,	Db	661	ATAACGCCAAGAGTGCCAAAGATGATGTTTTCGGGCCAAAGCGACGACGCGGACC	720
PT	feeding or reproduction of crop-damaging insects or insects of farm				
XX	animals.				
PS	Claim 2; Fig 3A-C; 119pp; English.	QY	721	GCCACGCGACGCCCGCCAGCTGATCCCGCTGTTTTCGGGTTTCATCGTCTGGGGCTG	780
XX		Db	721	GCCACGCGACGCCCGCCAGCTGATCCCGCTGTTTTCGGGTTTCATCGTCTGGGGCTG	780
CC	The invention describes an isolated tumour necrosis factor polypeptide	QY	781	GTGTTGCGCATTCCTCGCATTAACGATCTGGCAGACAAACGCTGATTCGATCTGGACAA	840
CC	(TNF). The polypeptide and polynucleotide are useful in controlling	Db	781	GTGTTGCGCATTCCTCGCATTAACGATCTGGCAGACAAACGCTGATTCGATCTGGACAA	840
CC	agriculturally important pests, particularly by modifying the growth,				
CC	feeding and/or reproduction of crop-damaging insects or insects of farm	QY	841	GAGCTGAAGAGCCTGAAGCGAGTCGTCGATAATCTCCAGCAGGCTTGGGCAATAACTAT	900
CC	animals. The polypeptide and polynucleotide are useful for modulating	Db	841	GAGCTGAAGAGCCTGAAGCGAGTCGTCGATAATCTCCAGCAGGCTTGGGCAATAACTAT	900
CC	epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus				
CC	the polypeptide and polynucleotide may be useful for treating,	QY	901	CTGGACGAGTTTCCGACGAGTTCCAAAGGAGTACGAGAATGCCCTCATCGACTATCCAAA	960
CC	ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal	Db	901	CTGGACGAGTTTCCGACGAGTTCCAAAGGAGTACGAGAATGCCCTCATCGACTATCCAAA	960
CC	dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-				
CC	like disorders, e.g. sparse hair, abnormal or missing teeth or sweat	QY	961	AAGGTGGATGGCTCACCAGATCAGGAGCAGCAGAGTGGCGATGGTCTGATTTCCATTT	1020
CC	gland aberrations in animals (e.g. insects and potentially humans),	Db	961	AAGGTGGATGGCTCACCAGATCAGGAGCAGCAGAGTGGCGATGGTCTGATTTCCATTT	1020
CC	endotoxic shock, inflammation, haemorrhagic necrosis of tumours,				
CC	cytotoxicity and obesity-linked insulin resistance, all of which involve	QY	1021	GCGGACGACGAGGACGACGACCTTAGCTATAGCTCTGTGGATGATGTTGGCGGACACTAC	1080
CC	TNF molecules. This sequence encodes the drosophila melanogaster tumour	Db	1021	GCGGACGACGAGGACGACGACCTTAGCTATAGCTCTGTGGATGATGTTGGCGGACACTAC	1080
CC	necrosis factor variant 2 (TNFV2) protein, described in the invention				
XX					
SQ	Sequence 2148 BP; 645 A; 510 C; 506 G; 487 T; 0 U; 0 Other;				
Query Match					
Best Local Similarity 100.0%; Score 2148; DB 6; Length 2148;					
Matches 2148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 GGCACGAGCGGAACGGAGCTTTAAAGTGAGAAAAAGAACCGGTAAATCAGAGATCCCAAG	60			
Db	1 GGCACGAGCGGAACGGAGCTTTAAAGTGAGAAAAAGAACCGGTAAATCAGAGATCCCAAG	60			
QY	61 CAAGCGCGTGGTCATGATGACGAGAAAGAAAGCTATCGGTTTCAGTTAACTACTTAC	120			
Db	61 CAAGCGCGTGGTCATGATGACGAGAAAGAAAGCTATCGGTTTCAGTTAACTACTTAC	120			
QY	121 CAAGATTGAATTTCCGCATCGGCAATTAATAAATAACATAAGTGCRACTCGTCCACT	180			
Db	121 CAAGATTGAATTTCCGCATCGGCAATTAATAAATAACATAAGTGCRACTCGTCCACT	180			
QY	181 GTGTGTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTGGTTTCGCTGTCCTTTATCGCAAA	240			
Db	181 GTGTGTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTGGTTTCGCTGTCCTTTATCGCAAA	240			
QY	241 TGATAAACTAGAAAAATCTTCGAAACTGTGTTTCGGCTGTCCTTTATCGCAAA	300			
Db	241 TGATAAACTAGAAAAATCTTCGAAACTGTGTTTCGGCTGTCCTTTATCGCAAA	300			
QY	301 GATCGGGAAGAGAAAAACAGCAGTAGACAAAACAAAGTGTGTAATAACAATCTGAAAAG	360			
Db	301 GATCGGGAAGAGAAAAACAGCAGTAGACAAAACAAAGTGTGTAATAACAATCTGAAAAG	360			
QY	361 GGCACCATCAGCAGCCGAGGGTTTATCTATATAGATTCGAGCTTATCATCTATGC	420			
Db	361 GGCACCATCAGCAGCCGAGGGTTTATCTATATAGATTCGAGCTTATCATCTATGC	420			
QY	421 TGTCTGTAGGTTGTTCTGTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTTATA	480			
Db	421 TGTCTGTAGGTTGTTCTGTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTTATA	480			
QY	481 TAAAGTGCGAACAAAGCTCGATTGGAAACAGCTGTTCGAGTGCCTTTGAGTGGTGGGCAAG	540			
Db	481 TAAAGTGCGAACAAAGCTCGATTGGAAACAGCTGTTCGAGTGCCTTTGAGTGGTGGGCAAG	540			



Db 1621 ATATGCTACAAACATCGCAGCACAGAAACGATTTATCGTCTTTCAAGGAGACACTCCA 1680  
Qy 1681 TTCCTGCAAGTCTTGAACACCGTCCCAACCAACATGCCACATAAGGTCACACCTGCCAC 1740  
Db 1691 TTCCTGCAAGTCTTGAACACCGTCCCAACCAACATGCCACATAAGGTCACACCTGCCAC 1740  
Qy 1741 ACGAGTGTCTGATCCCACTCTGGAACGAAACGAGAGGATCCATCTCTGAAGACATTCACAAC 1800  
Db 1741 ACGAGTGTCTGATCCCACTCTGGAACGAAACGAGAGGATCCATCTCTGAAGACATTCACAAC 1800  
Qy 1801 GATGCAATGAGTTCCTCGGAGGAGAAACCAACGAGACTCTTTGGCATCTTCAAGGTG 1860  
Db 1801 GATGCAATGAGTTCCTCGGAGGAGAAACCAACGAGACTCTTTGGCATCTTCAAGGTG 1860  
Qy 1861 TAAATTGAGAGATTTATCCCGGTGAGAGATGGAATACCAAGTTTAAGCTTTTGTCCCG 1920  
Db 1861 TAAATTGAGAGATTTATCCCGGTGAGAGATGGAATACCAAGTTTAAGCTTTTGTCCCG 1920  
Qy 1921 CGACTGCTCGTAATGCGATTATCGCCGAGCGGTGAATCCATTAGTTCTGTAGTACCTAGTC 1980  
Db 1921 CGACTGCTCGTAATGCGATTATCGCCGAGCGGTGAATCCATTAGTTCTGTAGTACCTAGTC 1980  
Qy 1981 TTAGTCACTCAAACTTAATCTCAATCGAATCGTGCAATCTGCATTAGTCAGAACGG 2040  
Db 1981 TTAGTCACTCAAACTTAATCTCAATCGAATCGTGCAATCTGCATTAGTCAGAACGG 2040  
Qy 2041 AGGAATATCATTTATTTTGTATATCTCTCGACTCTAAAGTGGAATATAAATAATA 2100  
Db 2041 AGGAATATCATTTATTTTGTATATCTCTCGACTCTAAAGTGGAATATAAATAATA 2100  
Qy 2101 TGTAGCTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCGAG 2148  
Db 2101 TGTAGCTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCGAG 2148

RESULT 2

ID ABK11678

XX ABK11678 standard; DNA; 2166 BP.

AC ABK11678;

XX 05-JUN-2002 (first entry)

XX DNA encoding tumour necrosis factor (TNF).

KW Tumour necrosis factor; TNF; pesticide; crop-damaging insect;  
KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;  
KW x-linked anhidrotic (hypohidrotic) ectodermal dysplasia;  
KW x-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;  
KW sparse hair; sweat gland aberration; ectotoxic shock; inflammation;  
KW haemorrhagic necrosis of tumour; cytotoxicity;  
KW obesity-linked insulin resistance; gene; ds.

OS Drosophila melanogaster.

XX Key Location/Qualifiers

FT misc\_difference 1..1878  
FT /tag= d  
FT /note= "Specifically claimed in claim 14"

FT CDS 652..1881

FT /tag= b

FT /product= "TNF"

FT /note= "Tumour necrosis factor"

FT sig\_peptide 652..807

FT /tag= a

FT mat\_peptide 808..1878

FT /tag= c

FT /label= mature\_TNF

XX US2002012968-A1.

PN 31-JAN-2002.

XX

PD

XX 20-MAR-2001; 2001US-00813329.  
XX PF  
XX 21-MAR-2000; 2000US-0190816P.  
XX PR  
XX (CARR/) CARROLL P M.  
XX (CHEN/) CHEN J.  
XX (RAMA/) RAMANATHAN C S.  
XX (XIAO/) XIAO H.  
XX (GUAN/) GUAN B.  
XX (BOWE/) BOWEN M A.  
XX Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;  
PI P-PSDB; AAU77716.  
XX WPI; 2002-195121/25.  
XX DR  
XX P-PSDB; AAU77716.  
XX  
XX New Drosophila tumor necrosis factor molecule, useful in controlling  
FT agriculturally important pests, e.g. comprises modifying the growth,  
FT feeding or reproduction of crop-damaging insects or insects of farm  
FT animals.  
XX  
XX Claim 2; Fig 1A-C; 119pp; English.  
XX  
XX The invention describes an isolated tumour necrosis factor polypeptide  
CC (TNF). The polypeptide and polynucleotide are useful in controlling  
CC agriculturally important pests, particularly by modifying the growth,  
CC feeding and/or reproduction of crop-damaging insects or insects of farm  
CC animals. The polypeptide and polynucleotide are useful for modulating  
CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus  
CC the polypeptide and polynucleotide may be useful for treating,  
CC ameliorating or preventing x-linked anhidrotic (hypohidrotic) ectodermal  
CC dysplasia and x-linked anhidrotic (hypohidrotic) ectodermal dysplasia-  
CC like disorders, e.g. sparse hair, abnormal or missing teeth or sweat  
CC gland aberrations in animals (e.g. insects and potentially humans),  
CC endotoxigenic shock, inflammation, haemorrhagic necrosis of tumours,  
CC cytotoxicity and obesity-linked insulin resistance, all of which involve  
CC TNF molecules. This sequence encodes the drosophila melanogaster tumour  
CC necrosis factor (TNF) protein, described in the invention  
XX  
XX Sequence 2166 BP; 641 A; 511 C; 520 G; 494 T; 0 U; 0 Other;  
SQ  
Query Match 98.4%; Score 2112.8; DB 6; Length 2166;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2126; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
Qy 1 GGCACGAGGCGAAGCGGACGTTTAAAGTGAGAAAAGAACCGGTAAATCAGAGATCCCAAG 60  
Db 19 GGCACGAGGCGAAGCGGACGTTTAAAGTGAGAAAAGAACCGGTAAATCAGAGATCCCAAG 78  
Qy 61 CAAGCGGTCGCTGCATGATAGCGAAGAAAAGGCTATCCGTTTCAGTTAACTACTTAC 120  
Db 79 CAAGCGGTCGCTGCATGATAGCGAAGAAAAGGCTATCCGTTTCAGTTAACTACTTAC 138  
Qy 121 CAAGATTGAATTTTCGCCATCGGCAAAATTAATAAATAATACATAAGTGCACACTCGTCCACT 180  
Db 139 CAAGATTGAATTTTCGCCATCGGCAAAATTAATAAATAATACATAAGTGCACACTCGTCCACT 198  
Qy 181 GTGTGTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTCGCTGTCGCTTTATCGCAACCAAGAAC 240  
Db 199 GTGTGTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTCGCTGTCGCTTTATCGCAACCAAGAAC 258  
Qy 241 TGATAAACTAGAAAATATCTTCGAAAACCTTGTTTTCGCGCTTTTCTTTTTCGTTAAATGCCC 300  
Db 259 TGATAAACTAGAAAATATCTTCGAAAACCTTGTTTTCGCGCTTTTCTTTTTCGTTAAATGCCC 318  
Qy 301 GATCGCGAAGAGAAAACCAAGCAGTAGACAAAACAAGTGTGGTAATACAAATCTGAAAAG 360  
Db 319 GATCGCGAAGAGAAAACCAAGCAGTAGACAAAACAAGTGTGGTAATACAAATCTGAAAAG 378  
Qy 361 GGCACCATCAGCAGCCCGAGGGGTTTATCTATATAGATGCGCAGCTTATCATCTCATGC 420  
Db 379 GGCACCATCAGCAGCCCGAGGGGTTTATCTATATAGATGCGCAGCTTATCATCTCATGC 438

QY	421	TGTCGTGAGGTTGTTCTGTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTCA	480
Db	439	TGTCGTGAGGTTGTTCTGTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTCA	498
QY	481	TAAAGTCGCAAAAGCTCGAATTGGAACAGCTCTCGAGTGCCTTTGAGTGGGTGGCAAG	540
Db	499	TAAAGTCGCAAAAGCTCGAATTGGAACAGCTCTCGAGTGCCTTTGAGTGGGTGGCAAG	558
QY	541	ATCGTCATCATCATCTCGTCGTCATTATCAACAGAAATCAGCATCAGCATCTGAGGCC	600
Db	559	ATCGTCATCATCATCTCGTCGTCATTATCAACAGAAATCAGCATCAGCATCTGAGGCC	618
QY	601	CGTTGCTTAAGATCCCCAGTTTTCATCAATTTATGACTGCCGAGACCCCTCAAGCCGTTT	660
Db	619	CGATGCTCTAAGATCCCCAGTTTTCATCAATTTATGACTGCCGAGACCCCTCAAGCCGTTT	678
QY	661	ATAACCCCAACGAGTCGCAACGATGATGTTTTTCGGCCAAAGCAGCAGCGCGACC	720
Db	679	ATAACCCCAACGAGTCGCAACGATGATGTTTTTCGGCCAAAGCAGCAGCGCGACC	738
QY	721	GCCCAGCGACGCAACCGCCAGCTGATCCCCCTGGTTTTGGGGTTTCATCGCTCTGGGGCTG	780
Db	739	GCCCAGCGACGCAACCGCCAGCTGATCCCCCTGGTTTTGGGGTTTCATCGCTCTGGGGCTG	798
QY	781	GTTCGTTGCCATTCTCGCATCTCGCATCTGGCAGACAACGCGTGTATCGCATCTGGACAAG	840
Db	799	GTTCGTTGCCATTCTCGCATCTCGCATCTGGCAGACAACGCGTGTATCGCATCTGGACAAG	858
QY	841	GAGCTGAAGAGCCTGAAGCGAGTCGTCGATTAATCTCCAGCAGCGTTTGGGATAAACTAT	900
Db	859	GAGCTGAAGAGCCTGAAGCGAGTCGTCGATTAATCTCCAGCAGCGTTTGGGATAAACTAT	918
QY	901	CTGGACGAGTTTCGACGAGTTTCCAAAAGAGTACGAGAATGCCCTCATCGACTATCCAAA	960
Db	919	CTGGACGAGTTTCGACGAGTTTCCAAAAGAGTACGAGAATGCCCTCATCGACTATCCAAA	978
QY	961	AAGGTGGATGGCCTCACCGATGAGGAGACGACGAGTGGGATGCGGATGCTTGGATTCATT	1020
Db	979	AAGGTGGATGGCCTCACCGATGAGGAGACGACGAGTGGGATGCGGATGCTTGGATTCATT	1038
QY	1021	GCGACGACGAGGACGACGAGCTTAGCTATAGCTCTGTGGATGATGTTGGCGCAGACTAC	1080
Db	1039	GCGACGACGAGGACGACGAGCTTAGCTATAGCTCTGTGGATGATGTTGGCGCAGACTAC	1098
QY	1081	GAGACTACACCGATATGTTAAATAACTCAACAAATGCATATCCGCGCACCGCCACA	1140
Db	1099	GAGACTACACCGATATGTTAAATAACTCAACAAATGCATATCCGCGCACCGCCACA	1158
QY	1141	TCTGAGACCACTGCTGAGGGCGAGGCGGAGACGACAGTGCATCCTCAGCGCTCAAAATGAT	1200
Db	1159	TCTGAGACCACTGCTGAGGGCGAGGCGGAGACGACAGTGCATCCTCAGCGCTCAAAATGAT	1218
QY	1201	GACAAATGTGTTGATGACTTTACAGCTCAATGCCCCAATAAAGAGCAGGAGAGAAA	1260
Db	1219	GACAAATGTGTTGATGACTTTACAGCTCAATGCCCCAATAAAGAGCAGGAGAGAAA	1278
QY	1261	TCTCGCTCGAATTGCCGATGTACGCAATGAGGAGCAGAAATTCACAGAAATCACACAG	1320
Db	1279	TCTCGCTCGAATTGCCGATGTACGCAATGAGGAGCAGAAATTCACAGAAATCACACAG	1338
QY	1321	CTTCAGGAAAGTCATCCAAATGAGGCAATTCCTCAAAAGAGAGCCCTGCACCTTCAACC	1380
Db	1339	CTTCAGGAAAGTCATCCAAATGAGGCAATTCCTCAAAAGAGAGCCCTGCACCTTCAACC	1398
QY	1381	CGTCGCGAAGTCAATTCGCCCATTCGCCACTCTAGTCGCGAAAGCAGATCCGAGAC	1440
Db	1399	CGTCGCGAAGTCAATTCGCCCATTCGCCACTCTAGTCGCGAAAGCAGATCCGAGAC	1458
QY	1441	TCGAGGCCAGACCCCATTTCCATTGAGCAGAGCGCGCTCAACAAGAAAGTATGGC	1500
Db	1459	TCGAGGCCAGACCCCATTTCCATTGAGCAGAGCGCGCTCAACAAGAAAGTATGGC	1518

[illegible]

```
FT sig_peptide 1..156
FT FT /*tag= a
FT mat_peptide 157..1218
FT FT /*tag= c
FT FT /label= mature_TNFv1
XX
PN US2002012968-A1.
XX
PD 31-JAN-2002.
XX
PF 20-MAR-2001; 2001US-00813329.
XX
PR 21-MAR-2000; 2000US-0190816P.
XX
PA (CARR/) CARROLL P M.
PA (CHEN/) CHEN J.
PA (RAMA/) RAMANATHAN C S.
PA (XIAO/) XIAO H.
PA (GUAN/) GUAN B.
PA (BOWE/) BOWEN M A.
XX
PI Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;
XX WPI; 2002-195121/25.
XX
DR New Drosophila tumor necrosis factor molecule, useful in controlling
PT agriculturally important pests, e.g. comprises modifying the growth,
PT feeding or reproduction of crop-damaging insects or insects of farm
PT animals.
XX
PS Claim 2; Fig 2A-B; 119pp; English.
XX
CC The invention describes an isolated tumour necrosis factor polypeptide
CC (TNF). The polypeptide and polynucleotide are useful in controlling
CC agriculturally important pests, particularly by modifying the growth,
CC feeding and/or reproduction of crop-damaging insects or insects of farm
CC animals. The polypeptide and polynucleotide are useful for modulating
CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus
CC the polypeptide and polynucleotide may be useful for treating,
CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal
CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-
CC like disorders, e.g. sparse hair, abnormal or missing teeth or sweat
CC gland aberrations in animals (e.g. insects and potentially humans),
CC endotoxic shock, inflammation, haemorrhagic necrosis of tumours,
CC cytotoxicity and obesity-linked insulin resistance, all of which involve
CC TNF molecules. This sequence encodes the drosophila melanogaster tumour
CC necrosis factor variant 1 (TNFv1) protein, described in the invention
XX
SQ Sequence 1221 BP; 349 A; 323 C; 312 G; 237 T; 0 U; 0 Other;

Query Match 53.2%; Score 1142.8; DB 6; Length 1221;
Best Local Similarity 96.3%; Pred. No. 4e-249;
Matches 1184; Conservative 0; Mismatches 37; Indels 9; Gaps 1;

QY 634 ATGACTGCGGAGACCTCAAGCCGTTTATACGCCAACGAGTGCCCAACGATGATGGTTT 693
DB 1 ATGACTGCGGAGACCTCAAGCCGTTTATACGCCAACGAGTGCCCAACGATGATGGTTT 60
QY 694 CCGGCCAAAGCGACGAGCGGCGACCGCCGACGACGACCGCCGACGATCCCGCTG 753
DB 61 CCGGCCAAAGCGACGAGCGGCGACCGCCGACGACGACCGCCGACGATCCCGCTG 120
QY 754 GTTTTGGGGTTTCATCGGTCTGGGGTGGTTCGTTTCGCTTCTCGCACTAACGATCTGGCAG 813
DB 121 GTTTTGGGGTTTCATCGGTCTGGGGTGGTTCGTTTCGCTTCTCGCACTAACGATCTGGCAG 180
QY 814 ACAACGCGTGTATCGCATCTGGACAAGGAGCTGAAGAGCCCTGAAGCGAGTCTGCGATAAT 873
DB 181 ACAACGCGTGTATCGCATCTGGACAAGGAGCTGAAGAGCCCTGAAGCGAGTCTGCGATAAT 240
QY 874 CTCGAGCAGCGTTTGGGCGATAACTATCTGAGCAGGTTTCGACGAGTTCCTCAAAAGGAGTAC 933
DB 241 CTCGAGCAGCGTTTGGGCGATAACTATCTGAGCAGGTTTCGACGAGTTCCTCAAAAGGAGTAC 300
```

## RESULT 4

ABL21473

ID ABL21473 standard; DNA; 978 BP.

XX ABL21473;

AC ABL21473;

XX ABL21473;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 15892.  
DE Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX Drosophila melanogaster.  
OS WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
XX 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX Claim 1; SEQ ID NO 15892; 21pp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB161176-AB130511), expressed DNA  
CC sequences (AB101840-AB161175) and the encoded proteins (AB57737-  
CC AB572072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 978 BP; 289 A; 261 C; 238 G; 190 T; 0 U; 0 Other;  
XX  
XX Query Match 42.2%; Score 907.4; DB 4; Length 978;  
XX Best Local Similarity 98.0%; Pred. No. 9.5e-196;  
XX Matches 936; Conservative 0; Mismatches 1; Indels 18; Gaps 1;  
XX  
QY 927 GGAGTACGAGAAATGCCCTCATCGACTATCCAAAAGGTGGATGGCTCAGGATGAGGA 986  
DB 24 GCAGTACGAGAAATGCCCTCATCGACTATCCAAAAGGTGGATGGCTCAGGATGAGGA 83  
QY 987 GGACGACGACGATGGCGATGGTCTGGATTCCATTGGCGACGACGAGGACGACGCTTAG 1046  
DB 84 GGACGACGACGATGGCGATGGTCTGGATTCCATTGGCGACGACGAGGACGACGCTTAG 143  
QY 1047 CTATAGCTCTGTGATGATGTGGCGGACGACTAGGAGCACTACCGGATATGTTAAATAA 1106  
DB 144 CTATAGCTCTGTGATGATGTGGCGGACGACTAGGAGCACTACCGGATATGTTAAATAA 203  
QY 1107 ACTCAACAAATGCACATACCGGACGACGACGACGACGACGACGACGACGACGACGACG 1166  
DB 204 ACTCAACAAATGCACATACCGGACGACGACGACGACGACGACGACGACGACGACGACG 263  
QY 1167 CGAGACGACGATGCTCCTCAGCCTCAATATGATGACAAATGTGTTCGATGACTTTACCAAG 1226  
DB 264 CGAGACGACGATGCTCCTCAGCCTCAATATGATGACAAATGTGTTCGATGACTTTACCAAG 323  
QY 1227 CTCAATGCCCAAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1286  
DB 324 CTCAATGCCCAAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 383  
QY 1287 TGAGGAGCAGAAATTTCAAGGAAATCACACAGAGCTTCAGGAGAAAGTCCAAATGAGGC 1346  
DB 384 TGAGGAGCAGAAATTTCAAGGAAATCACACAGAGCTTCAGGAGAAAGTCCAAATGAGGC 443

QY 1347 AACTTCCAAAGAGAGCCCTGCACCACTTCCACCGCTCGCAGAAATGCAATCCCGCCATCG 1406  
DB 444 AACTTCCAAAGAGAGCCCTGCACCACTTCCACCGCTCGCAGAAATGCAATCCCGCCATCG 503  
QY 1407 CCACCTCTAGTCCGCA-----AGCCAGATCCGAGAGACTCGAGGCC 1448  
DB 504 CCACCTCTAGTCCGCAAGGTGAATCTCTTCTTTCAGCCAGATCCGAGAGACTCGAGGCC 563  
QY 1449 AGCAGCCCAATTTCCACTTTGAGCAGAGCGGCGCTCACCAAGGAAGTATGGGCTACCATGG 1508  
DB 564 AGCAGCCCAATTTCCACTTTGAGCAGAGCGGCGCTCACCAAGGAAGTATGGGCTACCATGG 623  
QY 1509 AGATATGTACATAGGAAATGATAACGAGAGAAATCTTTATCAGGGACACTTTTCAAAACGCG 1568  
DB 624 AGATATGTACATAGGAAATGATAACGAGAGAAATCTTTATCAGGGACACTTTTCAAAACGCG 683  
QY 1569 CGATGGCGCTTTCAGCGTGACCAATACAGGCCTTATATAGTATACGCCCAAGATATGCTA 1628  
DB 684 CGATGGCGCTTTCAGCGTGACCAATACAGGCCTTATATAGTATACGCCCAAGATATGCTA 743  
QY 1629 CAACTTCCGACGACGACGAGAAATGATTTATCGTCTTTCAAGGAGACACTTCCATTTCTCTCA 1688  
DB 744 CAACTTCCGACGACGACGAGAAATGATTTATCGTCTTTCAAGGAGACACTTCCATTTCTCTCA 803  
QY 1689 GTGCTTGAACAGCGTGCGCCCAACCAATGCGCAATAGGTGCACTGCGCACACGAGTGG 1748  
DB 804 GTGCTTGAACAGCGTGCGCCCAACCAATGCGCAATAGGTGCACTGCGCACACGAGTGG 863  
QY 1749 TCTGATCCACCTTGAACGAGAAACGAGAGGATCCATCTGAGGACATTCACACGATCGCAA 1808  
DB 864 TCTGATCCACCTTGAACGAGAAACGAGAGGATCCATCTGAGGACATTCACACGATCGCAA 923  
QY 1809 TGCAGTCTTCGCGGAGGAGAAACACACGAGTACTTTTGGCATCTTCAAGGTGTAA 1863  
DB 924 TGCAGTCTTCGCGGAGGAGAAACACACGAGTACTTTTGGCATCTTCAAGGTGTAA 978  
XX  
XX RESULT 5  
XX ABL21472/c  
XX ID ABL21472 standard; DNA; 3324 BP.  
XX AC ABL21472;  
XX AC ABL21472;  
XX DT 26-MAR-2002 (first entry)  
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 15889.  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX Drosophila melanogaster.  
XX OS WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
XX 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX Claim 1; SEQ ID NO 15889; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention of  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3324 BP; 918 A; 715 C; 681 G; 1010 T; 0 U; 0 Other;  
 Query Match 36.3%; Score 779.4; DB 4; Length 3324;  
 Best Local Similarity 76.4%; Pred. No. 1.7e-166;  
 Matches 1184; Conservative 0; Mismatches 1; Indels 364; Gaps 4;  
 QY 927 GGAGTACGAGAAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCCTCACGGATGAGGA 986  
 Db 2301 GCAGTACGAGAAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCCTCACGGATGAGGA 2242  
 QY 987 GGACGACACGATGGCGATGCTCTGGATTCCATTCCGACGACGACGACGACGCTTAG 1046  
 Db 2241 GGACGACGATGGCGATGCTCTGGATTCCATTCCGACGACGACGACGACGCTTAG 2182  
 QY 1047 CTATAGCTCTGTGGATGATGTTGGCGCAGACTTACGAGGACTACACCGATATGTTAAATAA 1106  
 Db 2181 CTATAGCTCTGTGGATGATGTTGGCGCAGACTTACGAGGACTACACCGATATGTTAAATAA 2122  
 QY 1107 ACTCAACATGACATACCGGACCAACGCCACATCTGAGACCACTGCTGAGGCGAGGG 1166  
 Db 2121 ACTCAACATGACATACCGGACCAACGCCACATCTGAGACCACTGCTGAGGCGAGGG 2062  
 QY 1167 CGAGCGACAGTGCATCTCAGCCTCAAAATGATGACAAATGTTGCGATGATTTACGAG 1226  
 Db 2061 CGAGCGACAGTGCATCTCAGCCTCAAAATGATGACAAATGTTGCGATGATTTACGAG 2002  
 QY 1227 CTACAATGCCCAAAAAGAGCAGGAG----- 1254  
 Db 2001 CTACANTGCCCAAAAAGAGCAGGAGAGGTGAGCAGGATGAAATTTTCAGGGATTTA 1942  
 QY 1255 -----AGAAAATCTCGCTCGATTCGCGATGTACGC 1284  
 Db 1941 TCGTATTAAAGCCTCTGCTTTTCCCTCTTACAGAAAATCTCGCTCGATTCGCGATGTACGC 1882  
 QY 1285 AATGAGGAGCAGAAATATTCAAGGAAATCACACAGAGCTTCAGGAAAAGTCAATCCAAATGAG 1344  
 Db 1881 AATGAGGAGCAGAAATATTCAAGGAAATCACACAGAGCTTCAGGAAAAGTCAATCCAAATGAG 1822  
 QY 1345 GCAACTTCCAAAGAG----- 1359  
 Db 1821 GCAACTTCCAAAGAGAGGTAAAGTAGAACGTTCCGTTGAAAACAGGCAAGAAAACCATTA 1762  
 QY 1360 ----- 1359  
 Db 1761 ATGAAGTTATATGCTATGGTTAAAGATAGATAGTTAACTGGTAACCTGAACACATATG 1702  
 QY 1360 -----AGCCCTGCACCACTTCACCACCGTCGCAGAAATGCATTC 1397  
 Db 1701 TATACATTTTCTTTTACTTGAAGCCCTGCACCACTTCACCACCGTCGCAGAAATGCATTC 1642  
 QY 1398 CGCCATCGCCACCTCTCTAGTCCGCA----- 1424  
 Db 1641 CGCCATCGCCACCTCTCTAGTCCGCAAGAGGTGTAGTGTAGAAAGCGAAACAATTGGC 1582  
 QY 1425 ----- 1424  
 Db 1581 AAAAATCGAATACCAAAATCTCTATTCCGACCTCTGTTAAATTTGCGAGGTGAATCTCTTCTT 1522  
 QY 1425 ---AGCCAGATCGAGGAATCGAGGCCAGAGCCCAATTTCCACTTTCAGCAGGCGCGGT 1482  
 Db 1521 TCAGCCAGATCCGAGGAATCGAGGCCAGAGCCCAATTTCCACTTTCAGCAGGCGCGGT 1462

QY 1483 CACCAAGGAAGATG----- 1497  
 Db 1461 CACCAAGGAAGATGCGGTCAAGCGTGTACATACATACTACTAC 1402  
 QY 1498 -----GGCTACCATGGAGATATGTACATATG 1522  
 Db 1401 AGGCAACCTAATGAAGTCACTTCTTGAATTTCTCAGGCTACCATGGAGATATGTACATATG 1342  
 QY 1523 GAAATGATAACGAGAGAAACTCTTATCAGGAGACACTTTCAACGCGGATGGCTCTTGA 1582  
 Db 1341 GAAATGATAACGAGAGAAACTCTTATCAGGAGACACTTTCAACGCGGATGGCTCTTGA 1282  
 QY 1583 CGGTGACCAATACAGGCTATATTTACGTATATACGCCACAGATATGTCACAACTTCGCACG 1642  
 Db 1281 CGGTGACCAATACAGGCTATATTTACGTATATACGCCACAGATATGTCACAACTTCGCACG 1222  
 QY 1643 ACAGAAACGGAATTTATCGTCTTCAAGGAGACACTTCATCTCTGCACTGCTTGAACAGG 1702  
 Db 1221 ACCAGAAACGGAATTTATCGTCTTCAAGGAGACACTTCATCTCTGCACTGCTTGAACAGG 1162  
 QY 1703 TGCCCAACCAATGACCAATGAAGTGCACACCTGCGCACAGAGTGTCTGATCCACCTGG 1762  
 Db 1161 TGCCCAACCAATGACCAATGAAGTGCACACCTGCGCACAGAGTGTCTGATCCACCTGG 1102  
 QY 1763 AACGAAACGAGAGGATCCATCTGAAAGGACATTCACAAACGATCGCAATGCGAGTTCTGCGGG 1822  
 Db 1101 AACGAAACGAGAGGATCCATCTGAAAGGACATTCACAAACGATCGCAATGCGAGTTCTGCGGG 1042  
 QY 1823 AGGAAACAAACCGAAGTACTTTTGGCATCTTCAAGGTGTAATTTGGAGAGATTAATCCCG 1882  
 Db 1041 AGGAAACAAACCGAAGTACTTTTGGCATCTTCAAGGTGTAATTTGGAGAGATTAATCCCG 982  
 QY 1883 GTCAGAGATGGATACCACTTTAAAGCTTTTGTCCCGGACCTGCTCGTGAATCGGATTC 1942  
 Db 981 GTCAGAGATGGATACCACTTTAAAGCTTTTGTCCCGGACCTGCTCGTGAATCGGATTC 922  
 QY 1943 ATCGCCAGCGTGAATCCATCTAGTTAGTACCTAGTCTTAGTCACTCCAAACCTAATCT 2002  
 Db 921 ATCGCCAGCGTGAATCCATCTAGTTAGTACCTAGTCTTAGTCACTCCAAACCTAATCT 862  
 QY 2003 CAATCGGAATCGTGCACTATGCAATTTAGTCAAGAGACGAGGAGAAATCATATTTATTTGT 2062  
 Db 861 CAATCGGAATCGTGCACTATGCAATTTAGTCAAGAGACGAGGAGAAATCATATTTATTTGT 802  
 QY 2063 ATATACTCGTTCGACTCTAAAAGTGAATAAAAATATATGTAGCTATTA 2111  
 Db 801 ATATACTCGTTCGACTCTAAAAGTGAATAAAAATATATGTAGCTATTA 753  
 RESULT 6  
 AAQ21833  
 ID AAQ21833 standard; DNA; 390 BP.  
 XX  
 AC AAQ21833;  
 XX  
 DT 08-JUN-1992 (first entry)  
 XX  
 DE Randomising oligonucleotide used in SPERT mRNA prepn.  
 XX  
 KW Systematic polypeptide evolution by reverse translation; SPERT;  
 XX ligand binding; ss.  
 XX  
 OS Synthetic.  
 XX  
 FN W09202536-A.  
 XX  
 PD 20-FEB-1992.  
 XX  
 PF 02-AUG-1990; 90US-00561968.  
 XX  
 PR 02-AUG-1990; 90US-00561968.  
 XX















```

FT      CDS      /product= "glycoprotein X"
FT      complement(93636..94127)
FT      /tag= 1
FT      /product= "interferon regulatory factor 4"
FT      complement(111931..112443)
FT      /tag= m
FT      /product= "capsid protein IV"
FT      complement(123808..127296)
FT      /tag= n
FT      /product= "immediate early protein"
XX
PN      W09804576-A1.
XX
XX      05-FEB-1998.
XX
XX      22-JUL-1997; 97WO-US013346.
XX
PR      25-JUL-1996; 96US-00686243.
PR      25-JUL-1996; 96US-00686349.
PR      25-JUL-1996; 96US-00686350.
PR      25-JUL-1996; 96US-00687253.
PR      05-SEP-1996; 96US-00708678.
PR      10-OCT-1996; 96US-00728323.
PR      13-NOV-1996; 96US-00747887.
PR      13-NOV-1996; 96US-00748640.
PR      29-NOV-1996; 96US-00757669.
XX
PA      (UYCO ) UNIV COLUMBIA NEW YORK.
XX
XX      Chang Y, Bohenzky RA, Russo JJ, Edelman IS, Moore PS;
PI      WPI; 1998-130615/12.
XX
XX      New nucleic acid encoding Kaposi's sarcoma associated herpes virus
PT      proteins - useful for, e.g. detecting levels of HHV8 in, and preparation
PT      of vaccines for treatment of, HIV patients.
XX
PS      Example 2; Page 135-203; 230pp; English.
XX
XX      This sequence represents the long unique region and terminal repeat of
CC      the Kaposi's sarcoma-associated herpes virus (KSHV). KSHV is also known
CC      as human herpes virus 8 (HHV8). This sequence contains the DNAs of the
CC      invention which encode KSHV polypeptides selected from: (a) viral
CC      macrophage inflammatory protein (MIP) II; (b) viral interleukin-6 (IL-6);
CC      (c) viral IRF 1; (d) complement-binding protein; glycoproteins B, M or L;
CC      (d) capsid protein IV encoded by ORF65; and (e) immediate early protein
CC      encoded by ORF73. Labelled probes for the nucleic acid, proteins encoded
CC      by it, and antibodies (Ab) specific for the proteins are useful for
CC      detecting HHV8, specifically for diagnosis of Kaposi's sarcoma, in body
CC      fluids or tissue samples. HHV8 infections can be treated with antisense
CC      or triplex forming molecules or agents that bind specifically to the
CC      protein. Ab may be used for prophylaxis or treatment of HHV8 infection,
CC      while the protein can be used in protective vaccines. Ab may also be used
CC      to differentiate between lymphomas, and HHV8 may be implicated in many
CC      other lymphoproliferative diseases such as lymphomas, leukaemia,
CC      splenomegaly and mycosis fungoides. Cells and animals containing the
CC      nucleic acid are useful for drug screening. HHV8-derived peptides can be
CC      used as targets for antiviral drugs, e.g. dihydrofolate reductase gene
CC      can be inhibited with methotrexate. These can also be used to determine
CC      the immune status of a patient infected with HIV. HHV8 derived protein
CC      viral MIP III may be used as an anti-inflammatory agent for, e.g.
CC      treating rheumatoid arthritis. This sequence is stated as containing 81
CC      open reading frames. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ      Sequence 137507 BP; 32579 A; 37795 C; 35758 G; 31375 T; 0 U; 0 Other;

```

```

Query Match      2.6%; Score 55.8; DB 2; Length 137507;
Best Local Similarity 50.6%; Pred.No. 0.12;
Matches 135; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY      828 GCATCTGGACAGAGCTGAAGCGAGTCGTGATATCTCCAGCAGCGTTT 887
DB      126310 GGATGATGAGGAGGAGGAGAGACAGATGATGAGGAGGAGGAGGATGACGAGGAGGATCA 126251

```

```

QY      888 GGCATAAACTATCTGGACGAGTTTCGACGAGTTCCAAAAGGAGTACGAGATGCCCTCAT 947
DB      126250 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGA 126191
QY      948 CGACTATCCAAAAAAGGTGGATGCGCTCAGCGATGAGGAGGAGGACGACGATGGCGATGG 1007
DB      126190 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAGGA 126131
QY      1008 TCTGGATTCCATTGCGGACGACGAGGAGGACGACGATTTAGCTATAGCTCTGTGGATGATGT 1067
DB      126130 CGAGGAGGAGGACGAGGAGGAGGACGAGGAGGAGGAGGAGGAGGATGACGATGATGA 126071
QY      1068 TGGCGCAGACTACGAGGACTTACACCGA 1094
DB      126070 GGACAATGAGGACGAGGAGGATGACCA 126044

```

Search completed: September 14, 2004, 20:17:25  
Job time : 1231 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	66.8	3.1	390	3	US-09-197-649-7	Sequence 7, Appli
2	55.8	2.6	3489	2	US-08-728-323A-1	Sequence 1, Appli
3	55.8	2.6	3489	4	US-09-298-568-1	Sequence 1, Appli
4	55.8	2.6	3489	2	US-09-410-399-1	Sequence 1, Appli
5	55.8	2.6	32207	2	US-08-770-379-20	Sequence 20, Appl
6	55.8	2.6	32207	3	US-08-757-669A-20	Sequence 20, Appl
7	55.8	2.6	32207	4	US-09-230-371A-20	Sequence 20, Appl
8	51.6	2.4	1037	4	US-09-489-847-112	Sequence 112, Appl
9	51.4	2.4	3527	2	US-08-909-965C-7	Sequence 7, Appli
10	50.6	2.4	299	4	US-09-621-976-10211	Sequence 10211, A
11	50.6	2.4	1522	3	US-09-413-574-1	Sequence 1, Appli
12	50.2	2.3	1485	4	US-09-372-422A-39	Sequence 39, Appl
13	50.2	2.3	1736	3	US-09-182-816-22	Sequence 22, Appl
14	50.2	2.3	1736	3	US-09-182-816-24	Sequence 24, Appl
15	50.2	2.3	1736	3	US-09-471-528-22	Sequence 22, Appl
16	50.2	2.3	1736	3	US-09-471-528-24	Sequence 24, Appl
17	50.2	2.3	1736	3	US-09-634-530-22	Sequence 22, Appl
18	50.2	2.3	1736	3	US-09-634-530-24	Sequence 24, Appl
19	49.8	2.3	1133	4	US-09-916-204-1	Sequence 1, Appli
20	49.6	2.3	8100	4	US-09-554-337-4	Sequence 4, Appli
21	49.6	2.3	11517	1	US-07-920-281C-1	Sequence 1, Appli
22	49.6	2.3	11517	3	US-08-466-277-1	Sequence 1, Appli
23	49.6	2.3	15538	4	US-09-554-337-1	Sequence 1, Appli
24	49.4	2.3	441	4	US-09-601-537-10	Sequence 10, Appl
25	49.4	2.3	4121	4	US-09-601-537-9	Sequence 9, Appli
26	49.2	2.3	1637	4	US-09-205-258-178	Sequence 178, App
27	48.8	2.3	2610	4	US-09-545-814-1	Sequence 1, Appli





```
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-399-1

Query Match          2.6%; Score 55.8; DB 4; Length 3489;
Best Local Similarity 50.6%; Pred. No. 0.00063;
Matches 135; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy 828 GCATCTGACAAAGAGCTGAAGAGCCTGAAGCGAGTTCGCGATATCTCCAGCAGCGTTT 887
Db 987 GGATGATGAGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1046
Qy 888 GGGCATAAACTATCTGGACGAGTTCGACGAGTTCGACGAGTTCGACGAGTTCGACGAG 947
Db 1047 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGA 1106
Qy 948 CGACTATCCAAAAGGTGGATGGCTCAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1007
Db 1107 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAG 1166
Qy 1008 TCTGATTCCATTTCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1067
Db 1167 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1226
Qy 1068 TGGCGCAGACTACGAGGACTACACCGA 1094
Db 1227 GGACAATGAGGAGGAGGAGGATGACGA 1253

RESULT 5
US-08-770-379-20/c
; Sequence 20, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,379
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3489 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-770-379-20

Query Match          2.6%; Score 55.8; DB 2; Length 32207;
Best Local Similarity 50.6%; Pred. No. 0.0019;
Matches 135; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy 828 GCATCTGACAAAGAGCTGAAGAGCCTGAAGCGAGTTCGCGATATCTCCAGCAGCGTTT 887
Db 21010 GGATGATGAGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20951
Qy 888 GGGCATAAACTATCTGACGAGTTCGACGAGTTCGACGAGTTCGACGAGTTCGACGAG 947
Db 20950 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGA 20891
Qy 948 CGACTATCCAAAAGGTGGATGGCTCAGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1007
Db 20890 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAG 20831
Qy 1008 TCTGATTCCATTTCGCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1067
Db 20830 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 20771
Qy 1068 TGGCGCAGACTACGAGGACTACACCGA 1094
Db 20770 GGACAATGAGGAGGAGGATGACGA 20744

RESULT 6
US-08-757-669A-20/c
; Sequence 20, Application US/08757669A
; Patent No. 6183751
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,669A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-757-669A-20

Query Match
Best Local Similarity 2.6%; Score 55.8; DB 3; Length 32207;
Matches 135; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 828 GCATCTGGACAGGAGCTGAAGAGCCTGAAGCGAGTCGTGCGATATCTCCAGCAGCGTTT 887
Db 21010 GGATGATGAGGAGGACGAGGACAGACAGATGAGGAGGACGAGGAGGATGACGAGGAGGATGA 20951

QY 888 GGGCATAAACTATCTGGACAGCTTCGACGAGTTCCAAAAGGAGTACGAGATGCCCTCAT 947
Db 20950 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGA 20891

QY 948 CGACTATCCAAAAGGTGGATGGCTCACGGATGAGGAGGACGACGATGCGCGATGG 1007
Db 20890 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGA 20831

QY 1008 TCTGGATTCCATTTCGCGAGCGAGGACGACGCTTACGCTTAGCTCTGTGGATGATGT 1067
Db 20830 CGAGGAGGAGGACGAGGAGGAGGACGAGGAGGAGGACGAGGAGGATGACGATGATGA 20771

QY 1068 TGGCGCAGACTACGAGGACTACACCGA 1094
Db 20770 GGACAATGAGGACGAGGAGGATGACGA 20744

RESULT 7
US-09-230-371A-20/c
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20

Query Match
Best Local Similarity 2.6%; Score 55.8; DB 4; Length 32207;
Matches 135; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 828 GCATCTGGACAGGAGCTGAAGAGCCTGAAGCGAGTCGTGCGATATCTCCAGCAGCGTTT 887
Db 21010 GGATGATGAGGAGGACGAGGACAGACAGATGAGGAGGACGAGGAGGATGACGAGGAGGATGA 20951

QY 888 GGGCATAAACTATCTGGACAGCTTCGACGAGTTCCAAAAGGAGTACGAGATGCCCTCAT 947
Db 20950 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGA 20891

QY 948 CGACTATCCAAAAGGTGGATGGCTCACGGATGAGGAGGACGACGATGCGCGATGG 1007
Db 20890 CGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGA 20831

QY 1008 TCTGGATTCCATTTCGCGAGCGAGGACGACGCTTACGCTTAGCTCTGTGGATGATGT 1067
Db 20830 CGAGGAGGAGGACGAGGAGGAGGACGAGGAGGAGGACGAGGAGGATGACGATGATGA 20771

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-757-669A-20

Query Match
Best Local Similarity 2.4%; Score 51.6; DB 4; Length 1037;
Matches 75; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 2026 TTAGTCAGACGACGAGGAGGAAATCATATTTTGTATATATCTCGTTCGACTCTTAAAAA 2085
Db 896 TTGGGAGAGTTTGGAGGGGAGGGGATTTTTTTTAAANGATTTTNTTTTAAAAA 955

QY 2086 GTGATATAAATATATGTAGCTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2142
Db 956 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1012

RESULT 9
US-08-909-965C-7
; Sequence 7, Application US/08909965C
; Patent No. 5936078
; GENERAL INFORMATION:
; APPLICANT: Kuga Tetsuo
; APPLICANT: Nakagawa Satoshi
; APPLICANT: Sakaki Yoshiyuki
; APPLICANT: Zhao Nanding
; APPLICANT: Hashida Hideji
; TITLE OF INVENTION: NOVEL DNA, NOVEL POLYPEPTIDE
; AND NOVEL ANTIBODY
```

; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FITZPATRICK, CELLA, HARPER AND SCINTO  
; STREET: 277 Park Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10172-0194  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/909,965C  
; FILING DATE: August 12, 1997  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 322745/95  
; APPLICATION NUMBER: PCT/JP96/03630  
; FILING DATE: 12-No. 5936078-1995  
; FILING DATE: 12-Dec-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lawrence S. Perry  
; REGISTRATION NUMBER: 31865  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-2400  
; TELEFAX: 212-758-2982  
; TELEX: 236262  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3527 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; ORIGINAL SOURCE:  
; ORGANISM: human  
; IMMEDIATE SOURCE:  
; CLONE: F998  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 140 to 1084  
; IDENTIFICATION METHOD: by experiment  
US-08-909-965C-7

Query Match 2.4%; Score 51.4; DB 2; Length 3527;  
Best Local Similarity 69.3%; Pred. No. 0.0078;  
Matches 70; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
Qy 2042 GGAAATCATATTTATTTTGTATATCTCGTTCGACTCTAAAGTGAATAAAATATAT 2101  
Db 3419 GTAATAAAATCAATATGTATATAATCTCGTGAATCTAAAGTGAATAAAATATAT 3478  
Qy 2102 GTAGCTATTAAGTGAATAAAATCAATATGTATATAATCTCGTGAATCTAAAGTGAATAAAATATAT 2142  
Db 3479 AAAAAAATCAATATGTATATAATCTCGTGAATCTAAAGTGAATAAAATATAT 3519

RESULT 10  
US-09-621-976-10211  
; Sequence 10211, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm

; SEQ ID NO 10211  
; LENGTH: 299  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-10211

Query Match 2.4%; Score 50.6; DB 4; Length 299;  
Best Local Similarity 65.5%; Pred. No. 0.0036;  
Matches 74; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
Qy 2030 TCAGAACGCGAGGAAATCATATTTTGTATATCTCGTTCGACTCTAAAGTGA 2089  
Db 179 TTAAGACGAGTATGGGAACCAATTTTGTAAAGTAATATTTTCAACCTTTAAAAAAA 238  
Qy 2090 ATAAATATATGTAGCTATTAAAAAATCAATATTTTGTAAAGTAATATTTTCAACCTTTAAAAAAA 2142  
Db 239 AAAAAAATCAATATTTTGTAAAGTAATATTTTCAACCTTTAAAAAAA 291

RESULT 11  
US-09-413-574-1  
; Sequence 1, Application US/09413574  
; Patent No. 6235972  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; APPLICANT: Tagliani, Laura  
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof  
; FILE REFERENCE: 0964  
; CURRENT APPLICATION NUMBER: US/09/413,574  
; CURRENT FILING DATE: 1999-10-06  
; EARLIER APPLICATION NUMBER: 60/109,728  
; EARLIER FILING DATE: 1998-11-23  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1522  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (58) ... (1272)  
US-09-413-574-1

Query Match 2.4%; Score 50.6; DB 3; Length 1522;  
Best Local Similarity 57.1%; Pred. No. 0.0081;  
Matches 92; Conservative 0; Mismatches 59; Indels 0; Gaps 0;  
Qy 1982 TAGTCACTCCAAACCTAATCTCAATCGGAATCGTGCATATGTCATGTCAGAACGGA 2041  
Db 1360 TCGTCTATGCTGCTGCTGACTATCTTTTATTTCCATATATATTTTTCGGAATGCTTT 1419  
Qy 2042 GGAAATCATATTTATTTTGTATATCTCGTTCGACTCTAAAGTGAATAAAATATAT 2101  
Db 1420 CTAAGTCAATATTAATCAATATCAACGCTTACACCGTTACACCGTTAAAAA 1479  
Qy 2102 GTAGCTATTAAGTGAATAAAATCAATATGTATATAATCTCGTTCGACTCTAAAGTGAATAAAATATAT 2142  
Db 1480 AAAAAAATCAATATTTTGTAAAGTAATATTTTCAACCTTTAAAAAAA 1520

RESULT 12  
US-09-372-422A-39  
; Sequence 39, Application US/09372422A  
; Patent No. 6313375  
; GENERAL INFORMATION:  
; APPLICANT: Rudolf Jung  
; APPLICANT: Francois Barrieu  
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof  
; FILE REFERENCE: 0919  
; CURRENT APPLICATION NUMBER: US/09/372,422A  
; CURRENT FILING DATE: 1999-08-11  
; PRIOR APPLICATION NUMBER: US 60/098,692  
; PRIOR FILING DATE: 1998-08-31

```
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(1015)
US-09-372-422A-39

Query Match          2.3%; Score 50.2; DB 4; Length 1485;
Best Local Similarity 63.9%; Pred. No. 0.01;
Matches 76; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 2022 TGCATTAGTCAGAGACGAGGAGGAATCATATTTTGTATATCTCGTTCGACTCTA 2081
Db 1358 TGCATGTGTTGGAATCGGAAACACAAATATATATATGCGTTTCTTAAAAAAA 1417

QY 2082 AAAAGTGAATAAAATATATGTAGCTATTAAAAAAGTGAATAAAAGTGAATAAA 2140
Db 1418 AAAAAGTGAATAAAATATATGTAGCTATTAAAAAAGTGAATAAAAGTGAATAAA 1476

RESULT 13
US-09-182-816-22
; Sequence 22, Application US/09182816
; Patent No. 6143542
; GENERAL INFORMATION:
; APPLICANT: Wisniewski, Nancy
; APPLICANT: Silver, Gary M.
; APPLICANT: Lo, Katherine C.
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: NOVEL FLEA EPOXIDE HYDROLASE NUCLEIC ACID MOLECULES,
; FILE REFERENCE: FC-3-C1
; CURRENT APPLICATION NUMBER: US/09/182,816
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 08/989,510
; EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1736
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)..(1553)
US-09-182-816-22

Query Match          2.3%; Score 50.2; DB 3; Length 1736;
Best Local Similarity 68.0%; Pred. No. 0.011;
Matches 70; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 2040 GAGGAAATCATATTTTGTATATCTCGTTCGACTCTAAAGTGAATAAAATAT 2099
Db 1617 GAAATAAACGATATGGATTTTATTTCAAACTTGTCAAAATATAAAAAA 1676

QY 2100 ATGTAGCTATTAAAAAAGTGAATAAAAGTGAATAAAAGTGAATAAA 2142
Db 1677 AAAAAGTGAATAAAATATATGTAGCTATTAAAAAAGTGAATAAAAGTGAATAAA 1719

RESULT 14
US-09-182-816-24/c
; Sequence 24, Application US/09182816
; Patent No. 6143542
; GENERAL INFORMATION:
; APPLICANT: Wisniewski, Nancy
; APPLICANT: Silver, Gary M.
; APPLICANT: Lo, Katherine C.
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: NOVEL FLEA EPOXIDE HYDROLASE NUCLEIC ACID MOLECULES,
; FILE REFERENCE: FC-3-C1
; CURRENT APPLICATION NUMBER: US/09/182,816
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 08/989,510
; EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1736
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)..(1553)
US-09-182-816-24

Query Match          2.3%; Score 50.2; DB 3; Length 1736;
Best Local Similarity 68.0%; Pred. No. 0.011;
Matches 70; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 2040 GAGGAAATCATATTTTGTATATCTCGTTCGACTCTAAAGTGAATAAAATAT 2099
Db 1617 GAAATAAACGATATGGATTTTATTTCAAACTTGTCAAAATATAAAAAA 1676

QY 2100 ATGTAGCTATTAAAAAAGTGAATAAAAGTGAATAAAAGTGAATAAA 2142
Db 1677 AAAAAGTGAATAAAATATATGTAGCTATTAAAAAAGTGAATAAAAGTGAATAAA 1719

Search completed: September 15, 2004, 02:20:54
Job time : 181 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2004, 00:00:21 ; Search time 4276 Seconds  
(without alignments)  
2527.528 Million cell updates/sec

Title: US-09-813-329-5  
Perfect score: 2148  
Sequence: 1 ggcacgaggcgacgcgct.....aaaaaaaaaactcgag 2148

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2148	100.0		2148	9	US-09-813-329-5	Sequence 5, Appli
2	2112.8	98.4		2166	9	US-09-813-329-1	Sequence 1, Appli
3	1142.8	53.2		1221	9	US-09-813-329-3	Sequence 3, Appli
4	66.8	3.1		390	9	US-09-790-399-7	Sequence 7, Appli
5	57.2	2.7		966	13	US-10-425-114-7734	Sequence 7734, Ap
6	57.2	2.7		1269	13	US-10-424-599-46034	Sequence 46034, A
7	55.8	2.6		3489	13	US-09-894-273-1	Sequence 1, Appli
8	55.8	2.6		3489	15	US-10-294-804-1	Sequence 1, Appli
c	9	54.4	2.5	5306	15	US-10-311-455-483	Sequence 483, App
10	53.6	2.5		1915	15	US-10-225-567A-606	Sequence 606, App
11	53.2	2.5		2190	13	US-10-424-599-9903	Sequence 9903, A
12	52.8	2.5		2190	13	US-10-424-599-82347	Sequence 82347, A
c	13	52.8	2.5	372	13	US-10-424-599-38982	Sequence 38982, A
14	52.6	2.4		1898	16	US-10-264-049-287	Sequence 287, App

c	15	52.4	2.4	351	10	US-09-814-353-17579	Sequence 17579, A
	16	52.4	2.4	479	10	US-09-918-995-9339	Sequence 9339, Ap
	17	52.4	2.4	553	13	US-10-424-599-6106	Sequence 6106, Ap
	18	52.4	2.4	1506	17	US-10-437-963-40658	Sequence 40658, A
c	19	52.4	2.4	5858	15	US-10-311-455-1285	Sequence 1285, Ap
	20	52	2.4	399	13	US-10-424-599-26900	Sequence 26900, A
	21	51.8	2.4	3530	17	US-10-437-963-24712	Sequence 24712, A
c	22	51.6	2.4	279	13	US-10-424-599-93967	Sequence 93967, A
	23	51.6	2.4	417	13	US-10-424-599-116955	Sequence 116955, A
c	24	51.6	2.4	550	10	US-09-918-995-31149	Sequence 31149, A
c	25	51.6	2.4	603	17	US-10-021-323-5966	Sequence 5966, Ap
	26	51.6	2.4	672	15	US-10-198-846-8818	Sequence 8818, Ap
c	27	51.6	2.4	1029	17	US-10-767-701-11693	Sequence 11693, A
	28	51.6	2.4	1037	13	US-10-351-334-112	Sequence 112, App
c	29	51.6	2.4	1309	15	US-10-106-698-2117	Sequence 2117, Ap
	30	51.4	2.4	337	9	US-09-960-352-6976	Sequence 6976, Ap
c	31	51.4	2.4	1333	13	US-10-424-599-59340	Sequence 59340, A
	32	51.4	2.4	1933	13	US-10-424-599-33751	Sequence 33751, A
c	33	51.2	2.4	302	13	US-10-424-599-17202	Sequence 17202, A
	34	51.2	2.4	2122	16	US-10-374-780A-884	Sequence 884, App
c	35	51.2	2.4	14006	15	US-10-311-455-1931	Sequence 1931, Ap
	36	51	2.4	375	9	US-09-960-352-15014	Sequence 15014, A
c	37	51	2.4	418	13	US-10-424-599-118975	Sequence 118975, A
	38	51	2.4	472	17	US-10-437-963-57216	Sequence 57216, A
c	39	51	2.4	493	13	US-10-424-599-95204	Sequence 95204, A
	40	51	2.4	517	17	US-10-021-323-11054	Sequence 11054, A
c	41	51	2.4	527	9	US-09-864-761-16481	Sequence 16481, A
	42	51	2.4	622	13	US-10-424-599-3380	Sequence 3380, Ap
c	43	51	2.4	3510	10	US-09-766-511B-31	Sequence 31, Appl
	44	51	2.4	3510	15	US-10-241-220-38	Sequence 38, Appl
c	45	51	2.4	3552	10	US-09-814-353-19891	Sequence 19891, A

ALIGNMENTS

RESULT 1

US-09-813-329-5  
; Sequence 5, Application US/09813329  
; Patent No. US20020012968A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Subb Company  
; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole  
; TITLE OF INVENTION: Variants Thereof  
; FILE REFERENCE: D0016.np  
; CURRENT APPLICATION NUMBER: US/09/813,329  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,816  
; PRIOR FILING DATE: 2000-03-21  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 2148  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (634)..(1860)  
US-09-813-329-5

Query Match	100.0%	Score 2148;	DB 9;	Length 2148;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2148;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GGCAGCGGCGAAGCGGACGTTTAAAGTGAGAGAAACCGGTAATACAGATCCCAAG	60	
Db	1	GGCAGCGGCGAAGCGGACGTTTAAAGTGAGAGAAACCGGTAATACAGATCCCAAG	60	
Qy	61	CAAGCGGCGTGGTGCATGATAGCGAGAGAAAAAGCTATCCGTTTCAGTTAACTACTTAC	120	
Db	61	CAAGCGGCGTGGTGCATGATAGCGAGAGAAAAAGCTATCCGTTTCAGTTAACTACTTAC	120	
Qy	121	CAAGATTGAATTTCCGCCATCGGCAAAATTACTAAAAATACATAGTCAACTCGTCCACT	180	



; CURRENT FILING DATE: 2001-03-20																					
; PRIOR APPLICATION NUMBER: 60/190,816																					
; FILING DATE: 2000-03-21																					
; NUMBER OF SEQ ID NOS: 65																					
; SOFTWARE: PatentIn version 3.0																					
; SEQ ID NO 1																					
; LENGTH: 2166																					
; TYPE: DNA																					
; ORGANISM: Drosophila melanogaster																					
; FEATURE:																					
; NAME/KEY: CDS																					
; LOCATION: (652)..(1878)																					
US-09-813-329-1																					
Query Match																					
Best Local Similarity 98.4%; Score 2112.8; DB 9; Length 2166;																					
Matches 2126; Conservative 0; Mismatches 22; Indels 0; Gaps 0;																					
Qy	1	GGC	ACG	AGG	CGA	ACG	AGT	TAA	AGT	GAG	AAA	GAAA	CCG	GTA	AAAT	CAG	AGAT	CCCA	AG	60	
Db	19	GGC	ACG	AGG	CGA	ACG	AGT	TAA	AGT	GAG	AAA	GAAA	CCG	GTA	AAAT	CAG	AGAT	CCCA	AG	78	
Qy	61	CA	AGC	CGT	CGT	GC	AT	CA	TAG	CGA	AGAA	AAAA	AGC	TAT	CCG	TTT	CAG	TTA	ACT	TAC	120
Db	79	CA	AGC	CGT	CGT	GC	AT	CA	TAG	CGA	AGAA	AAAA	AGC	TAT	CCG	TTT	CAG	TTA	ACT	TAC	138
Qy	121	CA	AG	ATT	GA	AT	TT	CG	CA	T	CGG	CA	AA	T	T	A	C	T	A	A	180
Db	139	CA	AG	ATT	GA	AT	TT	CG	CA	T	CGG	CA	AA	T	T	A	C	T	A	A	198
Qy	181	GT	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	240
Db	199	GT	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	258
Qy	241	TG	AT	AA	CT	AG	AA	AT	CT	TG	AA	CT	AG	AA	AT	CT	TG	AA	CT	300	
Db	259	TG	AT	AA	CT	AG	AA	AT	CT	TG	AA	CT	AG	AA	AT	CT	TG	AA	CT	318	
Qy	301	GAT	CG	GA	AG	AA	AA	CA	AG	CA	TAG	CA	AA	CA	AG	TG	TG	TA	TA	CA	360
Db	319	GAT	CG	GA	AG	AA	AA	CA	AG	CA	TAG	CA	AA	CA	AG	TG	TG	TA	TA	CA	378
Qy	361	GG	C	AC	CA	T	AG	C	AG	C	CG	G	G	G	T	T	A	T	A	T	420
Db	379	GG	C	AC	CA	T	AG	C	AG	C	CG	G	G	G	T	T	A	T	A	T	438
Qy	421	TG	T	GT	G	AG	GT	T	GT	G	T	G	T	G	T	G	T	A	T	A	480
Db	439	TG	T	GT	G	AG	GT	T	GT	G	T	G	T	G	T	G	T	A	T	A	498
Qy	481	TAA	AG	TG	CG	A	CA	AA	AG	CT	CG	AT	TG	GA	AA	CA	AG	CT	TG	540	
Db	499	TAA	AG	TG	CG	A	CA	AA	AG	CT	CG	AT	TG	GA	AA	CA	AG	CT	TG	558	
Qy	541	AT	CG	T	CA	T	CA	T	CA	T	CA	T	CA	T	CA	T	CA	T	CA	600	
Db	559	AT	CG	T	CA	T	CA	T	CA	T	CA	T	CA	T	CA	T	CA	T	CA	618	
Qy	601	CG	GT	T	CT	TA	AG	AT	CC	CG	AG	TT	TC	AT	TA	AG	AT	CC	CG	660	
Db	619	CG	GT	T	CT	TA	AG	AT	CC	CG	AG	TT	TC	AT	TA	AG	AT	CC	CG	678	
Qy	661	ATA	CG	CA	AG	CG	CA	AG	CG	CA	AG	CG	CA	AG	CG	CA	AG	CG	CA	720	
Db	679	ATA	CG	CA	AG	CG	CA	AG	CG	CA	AG	CG	CA	AG	CG	CA	AG	CG	CA	738	
Qy	721	G	C	C	A	G	C	A	G	C	C	A	G	C	C	A	G	C	C	780	
Db	739	G	C	C	A	G	C	A	G	C	C	A	G	C	C	A	G	C	C	798	
Qy	781	G	T	C	G	T	G	C	A	T	A	C	G	A	T	T	G	G	A	840	
Db	799	G	T	C	G	T	G	C	A	T	A	C	G	A	T	T	G	G	A	858	

Qy	841	G	A	G	C	T	A	A	G	C	C	T	G	A	A	G	C	G	A	G	900
Db	859	G	A	G	C	T	A	A	G	C	C	T	G	A	A	G	C	G	A	G	918
Qy	901	C	T	G	A	C	G	A	G	T	T	C	C	A	A	A	G	A	G	T	960
Db	919	C	T	G	A	C	G	A	G	T	T	C	C	A	A	A	G	A	G	T	978
Qy	961	A	A	G	T	G	A	T	G	C	C	T	C	A	C	G	A	T	G	G	1020
Db	979	A	A	G	T	G	A	T	G	C	C	T	C	A	C	G	A	T	G	G	1038
Qy	1021	G	C	G	A	C	G	A	G	A	C	G	A	C	G	A	C	G	A	C	1080
Db	1039	G	C	G	A	C	G	A	G	A	C	G	A	C	G	A	C	G	A	C	1098
Qy	1081	G	A	G	A	C	T	A	C	C	G	A	T	G	T	T	A	A	A	C	1140
Db	1099	G	A	G	A	C	T	A	C	C	G	A	T	G	T	T	A	A	A	C	1158
Qy	1141	T	C	T	G	A	G	A	C	C	T	G	A	G	G	C	G	A	G	A	1200
Db	1159	T	C	T	G	A	G	A	C	C	T	G	A	G	G	C	G	A	G	A	1218
Qy	1201	G	A	C	A	T	G	T	T	C	G	A	T	G	A	T	G	A	T	G	1260
Db	1219	G	A	C	A	T	G	T	T	C	G	A	T	G	A	T	G	A	T	G	1278
Qy	1261	T	C	T	G	C	T	G	A	T	G	C	A	T	G	C	A	T	G	C	1320
Db	1279	T	C	T	G	C	T	G	A	T	G	C	A	T	G	C	A	T	G	C	1338
Qy	1321	C	T	T	C	A	G	A	A	A	G	T	C	C	A	A	A	G	A	A	1380
Db	1339	C	T	T	C	A	G	A	A	A	G	T	C	C	A	A	A	G	A	A	1398
Qy	1381	C	G	T	C	G	A	A	T	G	C	A	T	C	C	C	C	A	T	C	1440
Db	1399	C	G	T	C	G	A	A	T	G	C	A	T	C	C	C	C	A	T	C	1458
Qy	1441	T	C	G	A	G	C	A	G	C	C	A	T	T	T	C	C	A	T	T	1500
Db	1459	T	C	G	A	G	C	A	G	C	C	A	T	T	T	C	C	A	T	T	1518
Qy	1501	T	A	C	C	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	1560
Db	1519	T	A	C	C	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	1578
Qy	1561	C	A	A	A	C	G	C	G	A	T	G	C	G	A	T	G	C	G	A	1620
Db	1579	C	A	A	A	C	G	C	G	A	T	G	C	G	A	T	G	C	G	A	1638
Qy	1621	A	T	A	T	G	T	A	C	A	A	C	T	G	C	A	A	T	A	T	1680
Db	1639	A	T	A	T	G	T	A	C	A	A	C	T	G	C	A	A	T	A	T	1698
Qy	1681	T	T	C	T	G	C	A	G	T	T	G	A	C	G	T	T	G	A	C	1740
Db	1699	T	T	C	T	G	C	A	G	T	T	G	A	C	G	T	T	G	A	C	1758
Qy	1741	A	C	G	A	T	G	T	A	C	A	A	C	T	G	G	A	A	T	A	1800
Db	1759	A	C	G	A	T	G	T	A	C	A	A	C	T	G	G	A	A	T	A	1818
Qy	1801	G	A	T	C	G	A	A	T	G	C	A	A	T	G	C	A	A	T	G	1860
Db	1819	G	A	T	C	G	A	A	T	G	C	A	A	T	G	C	A	A	T	G	1878
Qy	1861	T	A	A	T	T	G	A	G	A	T	T	A	A	T	T	A	A	T	T	1920
Db	1879	T	A	A	T	T	G	A	G	A	T	T	A	A	T	T	A	A	T	T	1938



```

1921 CGACTGCTCGGAATGCGATTCAATGCCAGCGTGAATCCATTAGTTCGTAGTACCTAGTC 1980
Db CGACTGCTCGTGAATGCGAATTCATGCCAGCGTGAATCCATTAGTTCGTAGTACCTAGTC 1998
Qy TTAGTCACTCCAAACCTAATCTCAATCGGAATCGTGATCTGCATTAGTTCAGAGACGG 2040
Db TTAGTCACTCCAAACCTAATCTCAATCGGAATCGTGATCTGCATTAGTTCAGAGACGG 2058
Qy AGGAAATCATATTTATTTCTATATCTCGTTCGACTCTAAAAAGTCAATAAAAAATATA 2100
Db AGGAAATCATATTTATTTCTATATCTCGTTCGACTCTAAAAAGTCAATAAAAAATATA 2118
Qy TGTAGCTATTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2148
Db TGTAGCTATTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2166

RESULT 3
US-09-813-329-3
; Sequence 3, Application US/09813329
; Patent No. US20020012968A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Suibb Company
; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole
; TITLE OF INVENTION: Variants thereof
; FILE REFERENCE: D0016.np
; CURRENT APPLICATION NUMBER: US/09/813,329
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1221
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1218)
US-09-813-329-3

```

QY	1054	TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACCGGATATGTTAAATAAACTCAAC	1113
Db	421	TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACCGGATATGTTAAATAAACTCAAC	480
QY	1114	AATGCACATACCGGCACACGCCCCACATCTGAGACCATCTGAGGGCGAGGGCGAGACG	1173
Db	481	AATGCACATACCGGCACACGCCCCACATCTGAGACCATCTGAGGGCGAGGGCGAGACG	540
QY	1174	GACAGTGCATCCTCAGCCTCAAAATGATGACAATGTGTTTCGATGACTTTTACCAGCTACAAT	1233
Db	541	GACAGTGCATCCTCAGCCTCAAAATGATGACAATGTGTTTCGATGACTTTTACCAGCTACAAT	600
QY	1234	GCCCAAAAAGACGAGGAGAGAAAATCTCGTTCGATTCGCCGATGACGCAATGAGGAG	1293
Db	601	GCCCAAAAAGACGAGGAGAGAAAATCTCGTTCGATTCGCCGATGACGCAATGAGGAG	660
QY	1294	CAGAAATTTCAAGGAAATCACACAGAGCTTCAGGAAAAGTCATCCATGAGGCAACTTCC	1353
Db	661	CAGAAATTTCAAGGAAATCACACAGAGCTTCAGGAAAAGTCATCCATGAGGCAACTTCC	720
QY	1354	AAAGAGGCCCTGACCACTTCCACCCTGCGAGAATGCAATCCCGCCATCGCCACCTC	1413
Db	721	AAAGAGGAAATGCAATCCCGCCATCGCCACCTCTAGTCCGAAAGTGAACTCTC	775
QY	1414	CTAGTCCGAAAGCCAGATCGAGGACTCGAGGCCAGCAGCCCATTTCCACTTGAGCAGC	1473
Db	776	-----TTCTTTTCAGGCAGATCGAGGACTCGAGGCCAGCAGCCCATTTCCACTTGAGCAGC	831
QY	1474	AGCGGGCGTCAACCAAGGAAGTATGGGCTACCATGGAGATATGTACATAGGAATGATAAC	1533
Db	832	AGCGGGCGTCAACCAAGGAAGTATGGGCTACCATGGAGATATGTACATAGGAATGATAAC	891
QY	1534	GAGAGAACTCTTATCAGGAGACATTTTCAAACGGCGGATGGCGTCTTTGACGGTGACCAAT	1593
Db	892	GAGAGAACTCTTATCAGGAGACATTTTCAAACGGCGGATGGCGTCTTTGACGGTGACCAAT	951
QY	1594	ACAGGCCTATTTACGTATACGCCAGATGCTACAACTCTCGACGACCCAGAACGGA	1653
Db	952	ACAGGCCTATTTACGTATACGCCAGATGCTACAACTCTCGACGACCCAGAACGGA	1011
QY	1654	TTTATCGTCTTTCAAGGAGACATCCATTCTCGAGTGTCTTGAACCGGTGCCACCAAC	1713
Db	1012	TTTATCGTCTTTCAAGGAGACATCCATTCTCGAGTGTCTTGAACCGGTGCCACCAAC	1071
QY	1714	ATGCCATAGGTGCACACCTGCCACAGATGGTCTGATCCACCTGGAAAGAAACGAG	1773
Db	1072	ATGCCATAGGTGCACACCTGCCACAGATGGTCTGATCCACCTGGAAAGAAACGAG	1131
QY	1774	AGGATCCATCTGAAGGACATTTCAACACGATCGCAATGCGAGTCTCTCGGGGAGGAAACAAC	1833
Db	1132	AGGATCCATCTGAAGGACATTTCAACACGATCGCAATGCGAGTCTCTCGGGGAGGAAACAAC	1191
QY	1834	CGAAGTACTTTGGCATCTTCAAGGTGTAA	1863
Db	1192	CGAAGTACTTTGGCATCTTCAAGGTGTAA	1221

```

RESULT 4
US-09-790-399-7
; Sequence 7, Application US/09790399
; Patent No. US20020038000A1
; GENERAL INFORMATION:
; APPLICANT: Gold, Larry
; APPLICANT: Tuerk, Craig
; APPLICANT: Pribnow, David
; APPLICANT: Smith, Jonathan D.
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
; FILE REFERENCE: NEX02/C1-CON2
; CURRENT APPLICATION NUMBER: US/09/790,399
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 09/197,649
; PRIOR FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: 07/829,461

```

OTHER INFORMATION: Clone ID: 700727423\_FLI

DB 615 CGACGACGATGAAGACGAGAGAAAGTGAATGT 650





/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

/ FILE REFERENCE: PA133P1

/ CURRENT APPLICATION NUMBER: US/10/264,049

/ CURRENT FILING DATE: 2002-10-04

/ PRIOR APPLICATION NUMBER: PCT/US01/18569

/ PRIOR FILING DATE: 2001-06-07

/ PRIOR APPLICATION NUMBER: US 60/209,467

/ PRIOR FILING DATE: 2000-06-07

/ NUMBER OF SEQ ID NOS: 4360

/ SOFTWARE: PatentIn Ver. 3.1

/ SEQ ID NO 287

/ LENGTH: 1898

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: misc feature

/ LOCATION: (1895)..(1895)

/ OTHER INFORMATION: n equals a,t,g, or c

US-10-264-049-287

Query Match 2.4%; Score 52.6; DB 16; Length 1898;

Best Local Similarity 58.3%; Pred. No. 0.0069;

Matches 91; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1988 CTCCAAACTAATCTCAATCGGAATCGTGCTACTGCTAGTTCAGAAAGCGAGGAAAA 2047

DB 1742 CTCAGTCCCTACTCTGCTTGGGATAGTGTGAGCTTCATTTGTACACGTGTGACTTCGT 1801

QY 2048 TCATATTTATTTTGTATATATCTGCTTCGACTCTAAAAGTGAATAAATAATATGTAGCT 2107

DB 1802 CAGTTACAAACCCCAATAACTCTGTAGAGTGGAAAAA 1861

QY 2108 ATTA 2143

DB 1862 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1897

## RESULT 15

US-09-814-353-17579/c

/ Sequence 17579, Application US/09814353

/ Publication No. US20030165831A1

/ GENERAL INFORMATION:

/ APPLICANT: Lee, John

/ APPLICANT: Thompson, Pamela

/ APPLICANT: Lillie, James

/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND

/ TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

/ FILE REFERENCE: MRI-006B

/ CURRENT APPLICATION NUMBER: US/09/814,353

/ CURRENT FILING DATE: 2001-03-21

/ PRIOR APPLICATION NUMBER: US 60/191,031

/ PRIOR FILING DATE: 2000-03-21

/ PRIOR APPLICATION NUMBER: US 60/207,124

/ PRIOR FILING DATE: 2000-05-25

/ PRIOR APPLICATION NUMBER: US 60/211,940

/ PRIOR FILING DATE: 2000-06-15

/ PRIOR APPLICATION NUMBER: US 60/216,820

/ PRIOR FILING DATE: 2000-07-07

/ PRIOR APPLICATION NUMBER: US 60/220,661

/ PRIOR FILING DATE: 2000-07-25

/ PRIOR APPLICATION NUMBER: US 60/257,672

/ PRIOR FILING DATE: 2000-12-21

/ NUMBER OF SEQ ID NOS: 22037

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 17579

/ LENGTH: 351

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-09-814-353-17579

Query Match 2.4%; Score 52.4; DB 10; Length 351;

Best Local Similarity 69.6%; Pred. No. 0.0023;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 18:37:56 ; Search time 8239 Seconds  
(without alignments)  
7785.406 Million cell updates/sec

Title: US-09-813-329-5  
Perfect score: 2148  
Sequence: 1 ggcacggcgacgacgt.....aaaaaaaaaaactcgag 2148

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_estfun:\*  
15: em\_estom:\*  
16: em\_estom:\*  
17: em\_ges\_hum:\*  
18: em\_ges\_inv:\*  
19: em\_ges\_pln:\*  
20: em\_ges\_vrt:\*  
21: em\_ges\_fun:\*  
22: em\_ges\_nam:\*  
23: em\_ges\_mus:\*  
24: em\_ges\_pro:\*  
25: em\_ges\_rod:\*  
26: em\_ges\_phg:\*  
27: em\_ges\_vrl:\*  
28: gb\_ges1:\*  
29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	686.4	32.0	699	12 BI141711	BI141711 SD15611.5
2	678	31.6	686	12 BG636194	BG636194 SD13795.5
3	673.4	31.4	686	12 BI637732	BI637732 SD19794.5
4	666.8	31.0	688	12 BI638827	BI638827 SD21218.5

5	632.8	29.5	649	12	BG640917	BG640917 SD12092.5
6	599.2	27.9	604	9	AI260099	AI260099 LP03784.5
7	545.4	25.4	565	12	BI620499	BI620499 RH51659.5
8	500	23.3	508	14	CK135985	CK135985 SD18286.3
9	500	23.3	532	14	CK135525	CK135525 RH51659.3
10	477.4	22.2	495	12	BI636564	BI636564 SD18286.5
11	408.8	19.0	447	12	BI635992	BI635992 SD17545.5
12	129.8	6.0	612	13	BX560278	BX560278 BX560278
13	115.6	5.4	1101	29	CNS017WQ	AL108596 Drosophila
14	65.6	3.1	468	12	BG553005	BG553005 dab82609.
15	62.8	2.9	752	14	CD099026	CD099026 AGENCOURT
16	62.8	2.9	923	13	BU771970	BU771970 SJBELA02
17	62	2.9	697	13	BQ524571	BQ524571 NISC no05
18	61.8	2.9	935	14	CK017088	CK017088 AGENCOURT
19	61	2.8	887	13	BQ731479	BQ731479 AGENCOURT
20	59.6	2.8	1942	13	BU771792	BU771792 SJEJA08
21	59.4	2.8	1091	13	BU766419	BU766419 SJAERA02
22	59.2	2.8	863	14	CD752688	CD752688 AGENCOURT
23	59	2.7	866	12	BM015504	BM015504 603641716
24	59	2.7	887	12	BG166263	BG166263 602345444
25	58.6	2.7	2162	13	BU766490	BU766490 SJAEGA03
26	58.2	2.7	811	14	CF147545	CF147545 AGENCOURT
27	57.6	2.7	697	10	AW134038	AW134038 fl14903.y
28	57	2.7	198	14	CF209778	CF209778 CAB20005
29	57	2.7	299	14	CB366950	CB366950 OI305 oOk
30	57	2.7	664	14	CB066301	CB066301 PVBE12D11
31	57	2.7	795	28	AZ528485	AZ528485 ENTCM64TP
32	57	2.7	823	28	AZ676218	AZ676218 ENTKE36TR
33	57	2.7	828	13	BW249956	BW249956 BW249956
34	57	2.7	843	28	BH139532	BH139532 ENTNG88TF
35	57	2.7	1124	13	BQ421045	BQ421045 AGENCOURT
36	56.6	2.6	933	13	BX706375	BX706375 BX706375
37	56.4	2.6	180	12	BM307698	BM307698 eak32g05.
38	56.4	2.6	501	13	C84767	C84767 C84767 Dict
39	56.2	2.6	681	29	CNS02EOD	AL193990 Tetradon
40	56	2.6	645	12	BM165350	BM165350 EST567873
41	56	2.6	647	12	BM163120	BM163120 EST565643
42	56	2.6	671	12	BM160252	BM160252 EST562775
43	56	2.6	694	12	BM168242	BM168242 EST570765
44	56	2.6	712	12	BM394933	BM394933 50072-2-6
45	56	2.6	717	12	BM160500	BM160500 EST563023

ALIGNMENTS

RESULT 1  
BI141711  
LOCUS  
DEFINITION SD15611.5prime SD Drosophila melanogaster Schneider L2 cell culture  
EST 03-JUL-2001  
BI141711  
ACCSSION BI141711  
VERSION BI141711.1 GI:14594155  
KEYWORDS EST.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 699)  
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,  
Lewis,S. and Rubin,G.M.  
TITLE BGP/HMI Drosophila EST Project  
JOURNAL Unpublished (2001)  
COMMENT Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd. Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)  
hit genomic AE003831: arm:2R [4876906,5137815]  
estimated-cyto:46B13-46E4: 04/13/2001  
Plate: SD.156 row: A column: 11

High quality sequence stop: 698.

FEATURES

source  
1. .699  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="SD15611"  
/lab\_host="DH5-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell culture pot2"  
/note="Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 32.0%; Score 686.4; DB 12; Length 699;  
Best Local Similarity 99.7%; Pred. No. 1.6e-76;  
Matches 698; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 10 CGAAGCGAGCTTTAAAGTGAGAAAAAGAACCGGTAAATCAGAGATCCCAAGACGCGGT 69  
DB 1 CGAAGCGAGCTTTAAAGTGAGAAAAAGAACCGGTAAATCAGAGATCCCAAGACGCGGT 60  
QY 70 GCGTGCATGATAGCGAAGAAAAAGCTATCGTTTCAGTTAACTTACCTTACCAAGATGA 129  
DB 61 GCGTGCATGATAGCGAAGAAAAAGCTATCGTTTCAGTTAACTTACCTTACCAAGATGA 120  
QY 130 ATTTCCGCATCGGCAAAATTAATAAATACATAAGTCAACTGTCACGTGTGTGTGT 189  
DB 121 ATTTCCGCATCGGCAAAATTAATAAATACATAAGTCAACTGTCACGTGTGTGTGT 180  
QY 190 GTTTTTTTTTTTTTTTTTTTTTTTTTTTCGCTGTGCTTTTATCGCAAAACAAGAACTGATAAAC 249  
DB 181 G-TTTTTTTTTTTTTTTTTTTTTTTTTTTCGCTGTGCTTTTATCGCAAAACAAGAACTGATAAAC 239  
QY 250 TAGAAATATCTTGAGAACTGTGTTTCGCGCTTTTCTTTTGTGCTTAAATGCGGGA 309  
DB 240 TAGAAATATCTTGAGAACTGTGTTTCGCGCTTTTCTTTTGTGCTTAAATGCGGGA 299  
QY 310 AGAGAAAAACAAGCAGTAGACAAAACAAGTGTGTGTAATACAACTGAAAGGGCACCATC 369  
DB 300 AGAGAAAAACAAGCAGTAGACAAAACAAGTGTGTGTAATACAACTGAAAGGGCACCATC 359  
QY 370 AGCAGCCCGAGGGTTTATCTATAGATGTCGAGCTTATCATCTCATCTGTCTGTGA 429  
DB 360 AGCAGCCCGAGGGTTTATCTATAGATGTCGAGCTTATCATCTCATCTGTCTGTGA 419  
QY 430 GTTTGTTCTGTGCTCGTGTGTGTATCTTAAATACATAGATGTGTTTCAATATAAGTGG 489  
DB 420 GTTTGTTCTGTGCTCGTGTGTGTATCTTAAATACATAGATGTGTTTCAATATAAGTGG 479  
QY 490 ACAAAGCTCGATTGGAAACAGCTCGAGTGCCTTTCAGTGGGTGGGCAAGATCGTCATC 549  
DB 480 ACAAAGCTCGATTGGAAACAGCTCGAGTGCCTTTCAGTGGGTGGGCAAGATCGTCATC 539  
QY 550 ATCATATCTGTCGTCATTTACACAGAAATCAGCATCAGCATCGAGTGTGAGGCGCCGTTGCTC 609  
DB 540 ATCATATCTGTCGTCATTTACACAGAAATCAGCATCAGCATCGAGTGTGAGGCGCCGTTGCTC 599  
QY 610 TAAGATCCCAAGTGTTCATCAATTTATGACGCGGAGACCTCAAGCGCTTATTAAGCCCA 669  
DB 600 TAAGATCCCAAGTGTTCATCAATTTATGACGCGGAGACCTCAAGCGCTTATTAAGCCCA 659  
QY 670 ACAGTCCCAACGATGATGTTTTCGCGGCAAAAGCGACCA 709  
DB 660 ACAGTCCCAACGATGATGTTTTCGCGGCAAAAGCGACCA 699

RESULT 2

BG636194 686 bp mRNA linear EST 23-APR-2001  
LOCUS SD13795.5prime SD Drosophila melanogaster Schneider L2 cell culture  
DEFINITION pot2 Drosophila melanogaster cDNA clone SD13795 5 similar to

CG12919: FBan0012919 located on: 2R 46E1-46E1;.: 04/13/2001, mRNA sequence.  
BG636194  
BG636194.1 GI:13763731  
EST.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 686)  
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.  
BDGP/HHMI Drosophila EST Project  
Unpublished (2001)  
Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu  
hit genomic AB003831: arm:2R [4876906,5137815]  
estimated-cyto:46B13-46B4: 04/13/2001  
Plate: SD.137 row: H column: 11  
High quality sequence stop: 685.  
Location/Qualifiers  
1. .686  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="SD13795"  
/lab\_host="DH5-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell culture pot2"  
/note="Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 31.6%; Score 678; DB 12; Length 686;  
Best Local Similarity 99.3%; Pred. No. 1.8e-75;  
Matches 681; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1401 CCATCGCCACCTCTCTAGTCCGCAAGCCAGATCCGAGACTCGAGGCCAGAGCCCAATTT 1460  
DB 1 CCATCGCCACCTCTCTAGTCCGCAAGCCAGATCCGAGACTCGAGGCCAGAGCCCAATTT 60  
QY 1461 CCATTTGAGCAGCAGCGCGCTCACCAAGGAAGTATGGCTTACCATGGAGATATGTACAT 1520  
DB 61 CCACCTGAGCAGCAGCGCGCTCACCAAGGAAGTATGGCTTACCATGGAGATATGTACAT 120  
QY 1521 AGGAAATGATAACGAGAGAACTCTTATCAGGGAACATTTCAAACGCGCGATGCGCTTT 1580  
DB 121 AGGAAATGATAACGAGAGAACTCTTATCAGGGAACATTTCAAACGCGCGATGCGCTTT 180  
QY 1581 GACGCTGACCAATACAGGCTATATACGTATACGCCAGATATGCTACACACTTCGCA 1640  
DB 181 GACGCTGACCAATACAGGCTATATACGTATACGCCAGATATGCTACACACTTCGCA 240  
QY 1641 CGACAGAACGGAATTTATCGTCTTTCAAGGAGACACTTCCATTCCTCGAGTGTGTAACAC 1700  
DB 241 CGACAGAACGGAATTTATCGTCTTTCAAGGAGACACTTCCATTCCTCGAGTGTGTAACAC 300  
QY 1701 GGTGCCCAACCAATGCCATAGTGTGCACACTGCCACAGAGTGGTCTGATCCACT 1760  
DB 301 GGTGCCCAACCAATGCCATAGTGTGCACACTGCCACAGAGTGGTCTGATCCACT 360  
QY 1761 GGAAACGAAACGAGAGGATCCATCTGAAGGACATTTCAACAGCATCGCAATCTCTGCG 1820  
DB 361 GGAAACGAAACGAGAGGATCCATCTGAAGGACATTTCAACAGCATCGCAATCTCTGCG 420  
QY 1821 GGAGGAAACACCGAAGCTTACTTTGGCATCTTTCAAGGTGTAATTTGGAGAGATTTATCCC 1880  
|||||





/db xref="taxon:7227"  
/clone="SD121218"  
/lab\_host="DHS-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
culture pOT2"  
/notes="Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized  
fractionated cDNAs were directly ligated into pOT2.  
Plasmid cDNA library."

## ORIGIN

Query Match 31.0%; Score 666.8; DB 12; Length 688;  
Best Local Similarity 99.4%; Pred. No. 4.4e-74;  
Matches 680; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

```
QY 5 CGAGCGCAACGCGGTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64
Db |||
QY 7 CGAGCGCAACGCGGTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 66
Db |||
QY 65 CGCGTGCCTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTCAGTTAACTACTTACCAG 124
Db |||
QY 67 CGCGTGCCTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTCAGTTAACTACTTACCAG 126
Db |||
QY 125 ATTGAATTCGCCATCGGCAAAATTAATAAATACATAAGTGCACACTCGTCCACTGTGT 184
Db |||
QY 127 ATCGAATTCGCCATCGGCAAAATTAATAAATACATAAGTGCACACTCGTCCACTGTGT 186
Db |||
QY 185 GTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 244
Db |||
QY 187 GTTGTG--TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 244
Db |||
QY 245 AAACTAGAAATATCTTTGAGAAACTGTTTTCGCGCTTTCTTTTGTCTAATTGCCGATC 304
Db |||
QY 245 AAACTAGAAATATCTTTGAGAAACTGTTTTCGCGCTTTCTTTTGTCTAATTGCCGATC 304
Db |||
QY 305 GCGGAGAGAAACCAACAGCAGTAGACAAACCAAGTGTGTAATCAATCTGAAAGGGCA 364
Db |||
QY 305 GCGGAGAGAGAAACCAACAGCAGTAGACAAACCAAGTGTGTAATCAATCTGAAAGGGCA 364
Db |||
QY 365 CCATCAGCAGCCCGAGGGGTTTATCTATATAGATGTCGACGTTATCATCTCATGTGTC 424
Db |||
QY 365 CCATCAGCAGCCCGAGGGGTTTATCTATATAGATGTCGACGTTATCATCTCATGTGTC 424
Db |||
QY 425 TGTGAGTGTCTGTGTGCTCGTGTAGTATCTTAAATAATAGAGTGTTCATATAA 484
Db |||
QY 425 TGTGAGTGTCTGTGTGCTCGTGTAGTATCTTAAATAATAGAGTGTTCATATAA 484
Db |||
QY 485 GTGCGCAAAAGCTCGATTGGAACAGCTGTCGAGTGCCCTTGAGTGGGTGGGCAAGTCG 544
Db |||
QY 485 GTGCGCAAAAGCTCGATTGGAACAGCTGTCGAGTGCCCTTGAGTGGGTGGGCAAGTCG 544
Db |||
QY 545 TCATCATCATCATCGTCTCATTTATCAACAGAATCAGCATCAGCATCTGGAGGCCCGGT 604
Db |||
QY 545 TCATCATCATCATCGTCTCATTTATCAACAGAATCAGCATCAGCATCTGGAGGCCCGGT 604
Db |||
QY 605 TGCTCTAAGATCCCGAGTTCATCAATTATGATCGCGAGACCCCTCAAGCCGTTTATAA 664
Db |||
QY 605 TGCTCTAAGATCCCGAGTTCATCAATTATGATCGCGAGACCCCTCAAGCCGTTTATAA 664
Db |||
QY 665 CGCCCAACGAGTGGCAACGATGATG 688
Db |||
QY 665 CGCCCAACGAGTGGCAACGATGATG 688
Db |||
```

RESULT 5  
BG640917 649 bp mRNA linear EST 23-APR-2001  
LOCUS SD12092.5prime SD Drosophila melanogaster Schneider L2 cell culture  
DEFINITION pOT2 Drosophila melanogaster cDNA clone SD12092 5, mRNA sequence.  
ACCESSION BG640917  
VERSION BG640917.1 GI:13772843  
KEYWORDS EST.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 649)  
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
Lewis, S. and Rubin, G.M.  
BDGP/HHMI Drosophila EST Project  
Unpublished (2001)  
Contact: Stapleton, M.  
BDGP

Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
hit genomic AE003831: arm:2R [4876906,5137815]  
estimated-cyto:46B13-46B4: 04/13/2001  
Plate: SD.120 row: H column: 8  
High quality sequence stop: 597.  
Location/Qualifiers

FEATURES  
source

1..649  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="SD12092"  
/lab\_host="DHS-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
culture pOT2"  
/note="Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized  
fractionated cDNAs were directly ligated into pOT2.  
Plasmid cDNA library."

## ORIGIN

Query Match 29.5%; Score 632.8; DB 12; Length 649;  
Best Local Similarity 99.4%; Pred. No. 7.9e-70;  
Matches 646; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

```
QY 5 CGAGCGCAACGCGGTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64
Db |||
QY 2 CGAGCGCAACGCGGTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 61
Db |||
QY 65 CGCGTGCCTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTCAGTTAACTACTTACCAG 124
Db |||
QY 62 CGCGTGCCTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTCAGTTAACTACTTACCAG 121
Db |||
QY 125 ATTGAATTCGCCATCGGCAAAATTAATAAATACATAAGTGCACACTCGTCCACTGTGT 184
Db |||
QY 122 ATCGAATTCGCCATCGGCAAAATTAATAAATACATAAGTGCACACTCGTCCACTGTGT 181
Db |||
QY 185 GTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 244
Db |||
QY 182 GTTGTG--TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 239
Db |||
QY 245 AAACTAGAAATATCTTTGAGAAACTTGTGTTTTCGCGCTTTCTTTTGTCTAATTGCCGATC 304
Db |||
QY 240 AAACTAGAAATATCTTTGAGAAACTTGTGTTTTCGCGCTTTCTTTTGTCTAATTGCCGATC 299
Db |||
QY 305 GCGGAGAGAAACCAACAGCAGTAGACAAACCAAGTGTGTAATCAATCTGAAAGGGCA 364
Db |||
QY 300 GCGGAGAGAGAAACCAACAGCAGTAGACAAACCAAGTGTGTAATCAATCTGAAAGGGCA 359
Db |||
QY 365 CCATCAGCAGCCCGAGGGGTTTATCTATATAGATGTCGACGTTATCATCTCATGTGTC 424
Db |||
QY 360 CCATCAGCAGCCCGAGGGGTTTATCTATATAGATGTCGACGTTATCATCTCATGTGTC 419
Db |||
QY 425 TGTGAGTGTCTGTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTTCATATAA 484
Db |||
QY 420 TGTGAGTGTCTGTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTTCATATAA 479
Db |||
QY 485 GTGCGCAAAAGCTCGATTGGAACAGCTGTCGAGTGCCCTTGAGTGGGTGGGCAAGTCG 544
Db |||
QY 480 GTGCGCAAAAGCTCGATTGGAACAGCTGTCGAGTGCCCTTGAGTGGGTGGGCAAGTCG 539
Db |||
QY 545 TCATCATCATCATCGTCTCATTTATCAACAGAATCAGCATCAGCATCTGGAGGCCCGGT 604
Db |||
```

```
Db      540 TCATCATCATCGTCGTCATTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGGT 599
      605 TGCTCTAAGATCCCGAGGTTCATCAATATGACTGCCGAGACCCCTCAAG 654
      600 TGCTCTAAGATCCCGAGGTTCATCAATATGACTGCCGAGACCCCTCAAG 649

RESULT 6
LOCUS   AI1260099               604 bp mRNA linear EST 19-APR-2001
DEFINITION LP03784.5prime LP Drosophila melanogaster larval-early pupal pot2
Drosophila melanogaster cDNA clone LP03784 5prime, mRNA sequence.
ACCESSION AI1260099
VERSION   1
KEYWORDS  EST.
SOURCE    Drosophila melanogaster (fruit fly)
ORGANISM  Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 604)
          Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
          Lewis,S. and Rubin,G.M.
          BDGP/HMI Drosophila EST Project
          Unpublished (2001)
          Contact: Stapleton, M.
          BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic sequence D505033
Plate: 37 row: G column: 12
High quality sequence stop: 421.
Location/Qualifiers
  1. .604
     /organism="Drosophila melanogaster"
     /mol_type="mRNA"
     /db_xref="taxon:7227"
     /clone="LP03784"
     /sex="male and female"
     /dev_stage="larvae-pupae"
     /lab_host="DHS-alpha"
     /clone_lib="LP Drosophila melanogaster larval-early pupal
     pot2"
     /note="Organ: whole body; Vector: pot2; Site 1: EcoRI;
     Site 2: XhoI; Sized fractionated cDNAs were directly
     ligated into pot2. Plasmid cDNA library.
     "

FEATURES
     source
     1. .604

ORIGIN
Query Match      27.9%; Score 599.2; DB 9; Length 604;
Best Local Similarity 99.5%; Pred. NO. 1.3e-65;
Matches 601; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      10 CGAACGGACGTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAGCGCGT 69
Db      1 CGAACGGACGTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAGCGCGT 60

Qy      70 GCGTGCAATGATAGCGAAGAAAAAAGCTATCCGTTTCAGTTAACTACTATTACCAAGATTGA 129
Db      61 GCGTGCAATGATAGCGAAGAAAAAAGCTATCCGTTTCAGTTAACTACTATTACCAAGATTGA 120

Qy      130 ATTTGCGCATCGGCAAAATTACTAAAATACATAAGTCAACTGTCACCTGTGTGTGT 189
Db      121 ATTTGCGCATCGGCAAAATTACTAAAATACATAAGTCAACTGTCACCTGTGTGTGT 180

Qy      190 GTTTTTTTTTTTTTTTTTTTTGGCTGTGCGCTTTTATCGCAACAAGAACTGATAAAAC 249
Db      181 GTTTTTTTTTTTTTTTTTTTTGGCTGTGCGCTTTTATCGCAACAAGAACTGATAAAAC 240

Qy      250 TAGAAAATATCTTGAGAAACTTGTTCGCGCTTTCTTTTCTTAATTTGCCGATCGCGGA 309
```

---

```
Db      241 TAGAAAATATCTTGAGAAACTTGTTCGCGCTTTCTTTTGTGTAATTTGCCGATCGCGGA 300
Qy      310 AGAGAAAACAAGCAGTAGACAAAACAAGTGTGTAACAATCTGAAAAGGGCACCATC 369
Db      301 AGAGAAAACAAGCAGTAGACAAAACAAGTGTGTAACAATCTGAAAAGGGCACCATC 360
Qy      370 AGCAGCCCGAGGGGTTTATCTATATAGATGTCGCAGCTTATCATCTCATGCTCTGTGA 429
Db      361 AGCAGCCCGAGGGGTTTATCTATATAGATGTCGCAGCTTATCATCTCATGCTCTGTGA 420
Qy      430 GGTGTGTTCTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTCAATAAAGTCG 489
Db      421 GGTGTGTTCTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTCAATAAAGTCG 480
Qy      490 ACAAAGCTCGATTGGAAACAGCTGTCGAGTGCCCTTCAGTGGGTGGGCAAGATCGTCATC 549
Db      481 ACAAAGCTCGATTGGAAACAGCTGTCGAGTGCCCTTCAGTGGGTGGGCAAGATCGTCATC 540
Qy      550 ATCATCATCGTCGTCATTATCAACAGAAATCAGCATCAGCATCTCGAGGCCCGGTTGCTC 609
Db      541 ATCATCATCGCGCTCATTTATCAACAGAAATCAGCATCAGCATTTTGAGGCCCGGTTGCTC 600
Qy      610 TAAG 613
      601 TAAG 604

RESULT 7
LOCUS   BF1620499               565 bp mRNA linear EST 07-SEP-2001
DEFINITION RH51659.5prime RH Drosophila melanogaster normalized Head pFLC-1
Drosophila melanogaster cDNA clone RH51659 5, mRNA sequence.
ACCESSION BF1620499
VERSION   1
KEYWORDS  EST.
SOURCE    Drosophila melanogaster (fruit fly)
ORGANISM  Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 565)
          Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,
          Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,
          George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,
          Misra,S., Mungall,C.J., Munoo,J., Pacleb,J., Paragas,V., Park,S.,
          Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and
          Rubin,G.M.
          BDGP/HMI RH Drosophila EST Project
          Unpublished (2001)
          Contact: Stapleton, M.
          BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003831: arm:2R [4876906,5137815]
estimated-cyto:46B13-46E4: 08/22/2001
Plate: RH.516 row: E column: 11
High quality sequence stop: 442.
Location/Qualifiers
  1. .565
     /organism="Drosophila melanogaster"
     /mol_type="mRNA"
     /db_xref="taxon:7227"
     /clone="RH51659"
     /sex="male and female"
     /dev_stage="Adult"
     /lab_host="DHS-alpha Tona"
     /clone_lib="RH Drosophila melanogaster normalized Head
     pFLC-1"
     /note="Organ: head; Vector: pFLC1; Site 1: XhoI; Site 2:
     BamHI; Library was kindly generated by Piero Carninci at
```

the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

ORIGIN									
Query Match 25.4%; Score 545.4; DB 12; Length 565;									
Best Local Similarity 99.8%; Pred. No. 6.8e-59;									
Matches 546; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
Qy	5	CGAGGCCAACGGACGTTTAAAGTCAGAAACCGGTAAATCAGAGATCCCAAGCAAG	64						
Db	19	CGAGGCCAACGGACGTTTAAAGTCAGAAACCGGTAAATCAGAGATCCCAAGCAAG	78						
Qy	65	CGCGTGCCTGCATCATAGACGGAAGAAAAAGACTATCCGTTTTCACTTAACTACTTTACCAAG	124						
Db	79	CGCGTGCCTGCATCATAGACGGAAGAAAAAGCTATCCGTTTCAGTTAACTACTTTACCAAG	138						
Qy	125	ATTGAATTTCCGCATCGGGCAAAATTCTAAATAATACATAAGTGCACCTCGTCCACTGTGT	184						
Db	139	ATTGAATTTCCGCATCGGGCAAAATTCTAAATAATACATAAGTGCACCTCGTCCACTGTGT	198						
Qy	185	GTTCGTGTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTCGCTGTCCTTTATTCGCAAAACAAGACTGAT	244						
Db	199	GTTCGTGTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTCGCTGTCCTTTATTCGCAAAACAAGACTGAT	258						
Qy	245	AAAACTAGAAAAATCTTTGAAACCTTTGTTTTTCGCGCTTTTCTTTTTCCTAAATGCCGATC	304						
Db	259	AAAACTAGAAAAATCTTTGAAACCTTTGTTTTTCGCGCTTTTCTTTTTCCTAAATGCCGATC	318						
Qy	305	GCGGAGAGAAAAACAAGCAGTAGACAAAAACAAGTGGTGAATACATCTGAAAAAGGCCA	364						
Db	319	GCGGAGAGAAAAACAAGCAGTAGACAAAAACAAGTGGTGAATACATCTGAAAAAGGCCA	378						
Qy	365	CCATCAGACGCCGAGGGGTTTACTATAGATGTCGCACTTATCATCTCATGTGTC	424						
Db	379	CCATCAGACGCCGAGGGGTTTACTATAGATGTCGCACTTATCATCTCATGTGTC	438						
Qy	425	TGTGAGGTTGTTCTGTGTGTCGCTGTAGTATCTTAAATACATAGTGTGTTTCATATAAA	484						
Db	439	TGTGAGGTTGTTCTGTGTGTCGCTGTAGTATCTTAAATACATAGTGTGTTTCATATAAA	498						
Qy	485	GTGCGACAAAGCTCGATTGGAACAAGCTGTCGAGTGCCTTTGAGTGGTGGGCAAGATCG	544						
Db	499	GTGCGACAAAGCTCGATTGGAACAAGCTGTCGAGTGCCTTTGAGTGGTGGGCAAGATCG	558						
Qy	545	TCATCAT	551						
Db	559	TCATCAT	565						

RESULT	8
CKI35985	
LOCUS	508 bp mRNA linear EST 02-DEC-2003
DEFINITION	SD18286.3prime SD Drosophila melanogaster Schneider L2 cell culture
POT2	Drosophila melanogaster cDNA clone SD18286 3, mRNA sequence.
ACCESSION	CKI35985
VERSION	CKI35985.1 GI:38627921
KEYWORDS	EST.
SOURCE	Drosophila melanogaster (fruit fly)
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 508) Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S. and Rubin, G.M. BDGP/HMMI Drosophila EST Project Unpublished (2001) Other ESTs: SD18286.5prime Contact: Stapleton, M. BDGP
REFERENCE	
AUTHORS	Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA
TITLE	
JOURNAL	
COMMENT	Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting poly-T sequence has been removed. hit genomic AB003831: arm:2R [4876306,5137815]  
estimated-cyto:46B13-46E4: 02/08/2002  
Plate: SD.182 row: H column: 2  
High quality sequence stop: 456.  
Location/Qualifiers  
1. .508  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="SD18286"  
/lab\_host="PH5-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell culture pot2"  
/note="Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

F1846M2 CDS1A1 CDS1A1									
ORIGIN	Query Match	23.3%	Score 500;	DB 14;	Length 508;				
	Best Local Similarity	99.0%;	Pred. No. 3.3e-53;						
	Matches 503;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;				
QY	1603	TATTACGTATACGCCGAGATATGCTACAAACAACCTCGCACGACGAGAAACGGATTTATCGTC	1662						
DB	508	TATTACGTATACGCCCCAGATATGCTACAAACAACCTCGCACGACGAGAAACGGATTTATCGTC	449						
QY	1663	TTTCAAGGACACACTCCATTTCCTGCAGTGCTTGAACACGGTGCCCAACCAACATGCCACAT	1722						
DB	448	TTTCAAGGAGACACTCCATTTCCTGCAGTGCTTGAACACGGTGCCCAACCAACATGCCACAT	389						
QY	1723	AAGGTGCACACCTGCCACACGAGTGGTCTGATCCACTCGAAACGAAACGAGAGGATCCAT	1782						
DB	388	AAGGTGCACACCTGCCACACGAGTGGTCTGATCCACTCGAAACGAAACGAGAGATCCAT	329						
QY	1783	CTGAAGGACATTCACAAACGATCCGAATGCAATCTCTCGGGAGGGGAAACAAACCGAAGCTAC	1842						
DB	328	CTGAAGGACATTCACAAACGATCCGAATGCAATCTCTCGGGAGGGGAAACAAACCGAAGCTAC	269						
QY	1843	TTTGGCATCTTCAAGGTGTAAATTTGAGAGATTTATCCCGGTGAGAAAGATGGAATACCCAG	1902						
DB	268	TTTGGCATCTTCAAGGTGTAAATTTGAGAGATTTATCCCGGTGAGAAAGATGGAATACCCAG	209						
QY	1903	TTTAAAGCTTTTGTCCCGGACCTGCTCGTGAATGCGATTTCATCCCGAGGCTGAATCCATT	1962						
DB	208	TTTAAAGCTTTTGTCCCGGACCTGCTCGTGAATGCGATTTCATCCCGAGGCTGAATCCATT	149						
QY	1963	AGTTCGTAGTACCTAGTCTTAGTCACTCCAAACCTTAATCTCAATCGGAATCGTGCATACT	2022						
DB	148	AGTTCGTAGTACCTAGTCTTAGTCACTCAATAACTTAATCTCAATCGGAATCGTGCATACT	89						
QY	2023	GCATTAGTCAGAGACGGAGGAAATCATATTTATTTCTGTATATCTCGTTCGACTCTAA	2082						
DB	88	GCATTAGTCAGAGACGGAGGAAATCATATTTATTTCTGTATATCTCGTTCGACTCTAA	29						
QY	2083	AAAGTGAATAAAAAATATATGTAGCTATT	2110						
DB	28	AAAGTGAATAAAAAATATATGTAGCTATT	1						
RESULT 9	CK135525/c								
LOCUS	CK135525	532 bp	mRNA	linear	EST 02-DEC-2003				
DEFINITION	RH51659.3prime RH Drosophila melanogaster normalized Head pf1c-1								
	Drosophila melanogaster cDNA clone RH51659.3 similar to CG12919:								
	Fban0012919 GO:[] located on: 2R 46E1-46E1; 08/05/2002, mRNA								
	sequence.								
ACCESSION	CK135525								
VERSION	CK135525.1	GI:38627461							
KEYWORDS	EST.								

SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 532)  
AUTHORS Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Beriman, B.,  
Carlson, J., Chame, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E.,  
George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G.,  
Miera, S., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,  
Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and  
Rubin, G. M.  
TITLE BDGP/HMI RH Drosophila EST Project  
JOURNAL Unpublished (2001)  
COMMENT Other ESTs: RH51659.5prime  
Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
Based upon the presence of a XhoI site followed by a run of 14 or  
more T residues at the beginning of the sequence, this clone was  
polyadenylated. The resulting Poly-T sequence has been removed. hit  
genomic AE003831: arm:2R [4876906,5137815]  
estimated-cyto:46B13-46E4: 08/05/2002  
Plate: RH.516 row: E column: 11  
High quality sequence stop: 382.  
FEATURES  
source  
1..532  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="RH51659"  
/sex="male and female"  
/dev\_stage="Adult"  
/lab\_host="DHS-alpha Tona"  
/clone\_lib="RH Drosophila melanogaster normalized Head  
pFle-1"  
/note="Organ: head; Vector: pFle1; Site 1: XhoI; Site 2:  
BamHI; Library was kindly generated by Piero Carninci at  
the RIKEN. The library was normalized and excised using  
Cre recombinase. Plasmid cDNA library."  
ORIGIN  
Query Match 23.3%; Score 500; DB 14; Length 532;  
Best Local Similarity 99.0%; Pred. No. 3.2e-53;  
Matches 503; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1604 ATTAGTATAGCCCGAGATGCTACAACTCGCAGCAGCAACGGATTTATCGTCT 1663  
DB 532 ATTAGTATAGCCCGAGATGCTACAACTCGCAGCAGCAACGGATTTATCGTCT 473  
QY 1664 TTCAAGGAGACATCTCCATCTCTGAGTGCTTGAACACCGTGCCCAACCATGCCACATA 1723  
DB 472 TTCAAGGAGACATCTCCATCTCTGAGTGCTTGAACACCGTGCCCAACCATGCCACATA 413  
QY 1724 AGGTGCACCTGCCACAGAGTGCTGTGATCCACCTGGAACGAGAGGATCCATC 1783  
DB 412 AGGTGCACCTGCCACAGAGTGCTGTGATCCACCTGGAACGAGAGGATCCATC 353  
QY 1784 TGAAGGACATTCACAAAGATCGCAATGCAATGTCGGGAGGAAACAAACGAGCTACT 1843  
DB 352 TGAAGGACATTCACAAAGATCGCAATGTCGGGAGGAAACAAACGAGCTACT 293  
QY 1844 TTGGCATCTTCAAGGTGTAATTGGAGAGATTATCCCGGTGAGAAAGATGAATACCAAGT 1903  
DB 292 TTGGCATCTTCAAGGTGTAATTGGAGAGATTATCCCGGTGAGAAAGATGAATACCAAGT 233  
QY 1904 TTAGCTTTTGTCCCGGACTGCTCGTGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1963  
DB 232 TTAAGCTTTTGTCCCGGACTGCTCGTGAATGCAATGCAATGCAATGCAATGCAATGCAAT 173

QY 1964 GTTCGTAGTACTAGTCTTAGTCACTCAACACCTAATCTCAATCGGAATCGTGACTCTG 2023  
DB 172 GTTCGTAGTACTAGTCTTAGTCACTCAACACCTAATCTCAATCGGAATCGTGACTCTG 113  
QY 2024 CATTAGTCAGAGACGAGGAGAAATCATATTTATTTGTATATATATCTGTTGACTCTTAA 2083  
DB 112 CATTAGTCAGAGACGAGGAGAAATCATATTTATTTGTATATATATCTGTTGACTCTTAA 53  
QY 2084 AAGTGAATAAAAAATATATGTAGCTATTATTA 2111  
DB 52 AAGTGAATAAAAAATATATGTAGCTATTATTA 25  
RESULT 10  
BI636564  
LOCUS BI636564 495 bp mRNA linear EST 10-SEP-2001  
DEFINITION SD18286.Sprime SD Drosophila melanogaster Schneider L2 cell culture  
pOT2 Drosophila melanogaster cDNA clone SD18286 5 similar to  
CG12919: FBan0012919 GO:[] located on: 2R 46E1-46E1; 05/19/2001,  
mRNA sequence.  
ACCESSION BI636564  
VERSION BI636564.1 GI:15538774  
KEYWORDS EST.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 495)  
AUTHORS Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
Lewis, S. and Rubin, G. M.  
TITLE BDGP/HMI Drosophila EST Project  
JOURNAL Unpublished (2001)  
COMMENT Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
hit genomic AE003831: arm:2R [4876906,5137815]  
estimated-cyto:46B13-46E4: 05/19/2001  
Plate: SD.182 row: H column: 2  
High quality sequence stop: 491.  
FEATURES  
source  
1..495  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="SD18286"  
/lab\_host="DHS-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
culture pOT2"  
/note="vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized  
fractionated cDNAs were directly ligated into pOT2.  
Plasmid cDNA library."  
ORIGIN  
Query Match 22.2%; Score 477.4; DB 12; Length 495;  
Best Local Similarity 99.6%; Pred. No. 2.2e-50;  
Matches 489; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 931 TACGAGAAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCCCTCAGCGATGAGGAGAC 990  
DB 6 TACGAGAAATGCCCTCATCGACTATCC-CAAAAGGTGGATGGCCCTCAGCGATGAGGAGAC 64  
QY 991 GACACGATGCGATGCTGTGGATTCCATTGCGGACGACGAGGACGACGCTTAGCTAT 1050  
DB 65 GACGACGATGCGATGCTGTGGATTCCATTGCGGACGACGAGGACGACGCTTAGCTAT 124  
QY 1051 AGCTCTGTGATGATGTTGGCGGACGACTACGAGGACTACACCGATATGTTAAATAACTC 1110  
DB 125 AGCTCTGTGATGATGTTGGCGGACGACTACGAGGACTACACCGATATGTTAAATAACTC 184

```
QY 1111 AACAAATGCATACCGCACCGCCACATCTGAGACCACTGCTGAGGGCGAGGGCGAG 1170
|||||
Db 185 AACAAATGCATACCGGCACCGCCACATCTGAGACCACTGCTGAGGGCGAGGGCGAG 244
|||||
QY 1171 ACGGACAGTGCATCCTCAGGCTCAAAATGATGACAATGTGTTGATGACTTTTACCGCTAC 1230
|||||
Db 245 ACGGACAGTGCATCCTCAGGCTCAAAATGATGACAATGTGTTGATGACTTTTACCGCTAC 304
|||||
QY 1231 AATGCCACAAAAAGAGAGGAGAGAAAAATCTCGCTCGATGTCGGATGACGAAATGAG 1290
|||||
Db 305 AATGCCACAAAAAGAGAGGAGAGAAAAATCTCGCTCGATGTCGGATGACGAAATGAG 364
|||||
QY 1291 GAGCAGAAATTCAGGAATACACAGAGCTTCAGGAAAGTCATCCAAATGAGGCAACT 1350
|||||
Db 365 GAGCAGAAATTCAGGAATACACAGAGCTTCAGGAAAGTCATCCAAATGAGGCAACT 424
|||||
QY 1351 TCCAAAGAGAGCCCTGCACCACTTCACCACTCGCGAGAAATGCAATCCCGCCATCGCCAC 1410
|||||
Db 425 TCCAAAGAGAGCCCTGCACCACTTCACCACTCGCGAGAAATGCAATCCCGCCATCGCCAC 484
|||||
QY 1411 CTCTAGTCCG 1421
|||||
Db 485 CTCTAGTCCG 495
|||||

RESULT 11
LOCUS B1635992 447 bp mRNA linear EST 10-SEP-2001
DEFINITION SD17545.5prime SD Drosophila melanogaster Schneider L2 cell culture
pOT2 Drosophila melanogaster cDNA clone SD17545 5, mRNA sequence.
ACCESSION B1635992
VERSION B1635992.1 GI:15538202
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 447)
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Teang,G.,
Lewis,S. and Rubin,G.M.
TITLE BGP/HMI Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT Contact: Stapleton, M.
BGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003831: arm:2R [4876906,5137815]
estimated-cyto:46B13-46E4: 05/19/2001
Plate: SD 175 row: D column: 9
High quality sequence stop: 445.
FEATURES
Location/Qualifiers
1..447
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="SD17545"
/lab_host="DH5-alpha"
/clone_lib="SD Drosophila melanogaster Schneider L2 cell
culture pOT2"
/note="vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized
fractionated cDNAs were directly ligated into pOT2.
Plasmid cDNA library."

ORIGIN
Query Match 19.0%; Score 408.8; DB 12; Length 447;
Best Local Similarity 99.1%; Pred. No. 8.4e-42;
Matches 422; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 5 CGAGCGAAGCGAGCTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64
|||||
```

```
Db 24 CGAAGCGAAGCGAGCTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 83
|||||
QY 65 CGCGTCGTCGATGATAGCGAAGAAAAAGCTATCCGTTTCAGTTAACTACTTACCAAG 124
|||||
Db 84 CGCGTCGTCGATGATAGCGAAGAAAAAGCTATCCGTTTCAGTTAACTACTTACCAAG 143
|||||
QY 125 ATTGAATTTCCGCATCGGGCAAAATTACTAAAAATACATAAAGTGCAACTCGTCCACTGTGT 184
|||||
Db 144 ATCGAATTTCCGCATCGGGCAAAATTACTAAAAATACATAAAGTGCAACTCGTCCACTGTGT 203
|||||
QY 185 GTTGTGTTTTTTTTTTTTTTTTTTTGGTTTTTCGCTGTCCTTTATPCGCAAAACAAGAACTGAT 244
|||||
Db 204 GTTGTG- TTTTTTTTTTTTTTTTTTTGGTTTTTCGCTGTCCTTTATPCGCAAAACAAGAACTGAT 261
|||||
QY 245 AAAACTAGAAAAATCTTGAGAACTTGTTTTTCGCGCTTTTCTTTTGTCTAATTCGCCGATC 304
|||||
Db 262 AAAACTAGAAAAATCTTGAGAACTTGTTTTTCGCGCTTTTCTTTTGTCTAATTCGCCGATC 321
|||||
QY 305 GGGGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGGTAATACAACTCTGAAAAAGGCA 364
|||||
Db 322 GCGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGGTAATACAACTCTGAAAAAGGCA 381
|||||
QY 365 CCATCAGAGCCGAGGGGTTTTATCTATATAGATGTCGAGCTTATCATCTCATGTCTGC 424
|||||
Db 382 CCATCAGAGCCGAGGGGTTTTATCTATATAGATGTCGAGCTTATCATCTCATGTCTGC 441
|||||
QY 425 TGTGAG 430
|||||
Db 442 TGTGAG 447
|||||

RESULT 12
LOCUS B16360278/c 612 bp mRNA linear EST 10-OCT-2003
DEFINITION Glossina morsitans morsitans adult infected gut Glossina
morsitans morsitans cDNA clone Tse49a10_plc, mRNA sequence.
ACCESSION B16360278
VERSION B16360278.1 GI:33368472
KEYWORDS EST.
SOURCE Glossina morsitans morsitans
ORGANISM Glossina morsitans morsitans
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Hippoboscidae; Glossinidae; Glossina.
REFERENCE 1 (bases 1 to 612)
AUTHORS Lehane,M.J., Aksoy,S., Gibson,W., Kerhornou,A., Berriman,M.,
Hamilton,J., Soares,M.B., Bonaldo,M.F., Lehane,S. and Hall,N.
TITLE Adult midgut expressed sequence tags from the tsetse fly Glossina
morsitans morsitans and expression analysis of putative immune
response genes
JOURNAL Genome Biol. 4 (10), R63 (2003)
MEDLINE 22881942
PUBMED 14519198
COMMENT Contact: Hall N
Pathogen Sequencing Unit
The Sanger Institute The Wellcome Trust Genome Campus
Hinxton, Cambridge, CB10 1SA, UK
Request for clones, please contact: Mike Lehane
Prof. M.J.Lehane
School of Biological Sciences,
University of Wales,
Bangor LL57 2UW
All clones with suffix g1c are reverse primer reads starting at 5'
end of the cDNA all plc reads are from
the 3' end.
FEATURES
Location/Qualifiers
1..612
/organism="Glossina morsitans morsitans"
/mol_type="mRNA"
/sub_species="morsitans"
/db_xref="taxon:37546"
/clone="Tse49a10_plc"
/tissue_type="adult infected gut"
```

/clone\_lib="Glossina morsitans morsitans adult infected gut"  
/note="country: Zimbabwe; EST from adult gut infected with T.brucei"

```
ORIGIN
Query Match      6.0%; Score 129.8; DB 13; Length 612;
Best Local Similarity 64.5%; Pred. No. 3.9e-07;
Matches 194; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy 1571 ATGGCGCTTCACGGTGACCAATACAGCCCTATTAGCTATAGCCCGAGATATGCTACA 1630
Db 607 ATGGTGCTTCGACTATTCCGAATCGGGCTTTTATTATGTTTATGCTCAAGTATGTTATA 548

Qy 1631 ACAACTCCACGACGACCAAGGATTATCGTCTTTCAAGGAGACACTCCATTCCTGCGAGT 1690
Db 547 ATACACGACGATCATTAATGGTTTGTGTTATTCATGGTCAAGGACCCCTTTTACAAAT 488

Qy 1691 GCTTGAACACGGTGCCCAACCAATGCGACATAAGGTGCACACCTGCCACACGAGTGGTC 1750
Db 487 GCTTCAAACTGTACCCACTAACATGTCACCTGAAATCCACACTTGTCTACGAGCGGTT 428

Qy 1751 TGATCCACCTGGACGAAACGAGAGGATCCATCTGAAGGACATTCACAAACGATCGCAATG 1810
Db 427 TAATTTATTTAAAAACACATGAAATAATTTACATTTAAGAGATTTTTCACAGTGACCGCAATG 368

Qy 1811 CAGTTCTCGGGAGGGAACCAACGAGCTACTTTGGCATCTTCAAGGTGTAAATTCGAG 1870
Db 367 CGGTCTGAAGGATTCAAACAATAGAAGTTACTTTGGTTTAAAGAAATCTGAAAAAGAAA 308

Qy 1871 A 1871
Db 307 A 307

RESULT 13
CNS017WQ      1101 bp      DNA      linear      GSS 26-JUL-1999
LOCUS      Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION      BACN37D20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION      AL108596.1 GI:5628900
VERSION      GSS.
KEYWORDS      Drosophila melanogaster (fruit fly)
SOURCE      Drosophila melanogaster
ORGANISM      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1101)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT      - Web : www.genoscope.cns.fr
      Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -- This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES      Location/Qualifiers
source      1..1101
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACN37D20"
            /clone_lib="DrosBAC"
            /plasmid="pBelobAC11"
            /note="end : T7"

ORIGIN
Query Match      3.1%; Score 65.6; DB 12; Length 468;
Best Local Similarity 48.9%; Pred. No. 47;
Matches 176; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

Qy 859 CGAGTCTCGATATATCTCCAGCAGCGTTTCGGGATAAATCTCTGGACGAGTTCGACGAG 918
Db 21 CGTGACGATGATGATGACGATGATGACGATGATGATGATGATGATGATGATGATGATGAC 80
```

Qy	919	TTCCAAAGGAGTAGTACGAGATGCCCTCATCGACTATCCAAAAAGGTGGATGGCCTCAGC	978
Db	81	GATGATGACGACGATGATGATGATGATGACGACGATGATGAAGATGATGATGACGACGAC	140
Qy	979	GATCAGGAGGACGACGATGCGGATGGTCTGGATTCCATTGCGGACGACGAGGACGAC	1038
Db	141	GATGATGATAAGACGATGATGACGATGATGACGACGATGATGACGACGATGATGAC	200
Qy	1039	GACGTTAGCTATAGCTCTGTGGATGATGTTGGCGGACGACTACGAGGACTACACGATATG	1098
Db	201	GACGACGACGATGACACGATGATGATGATGATCTAAAAAGGATGACGACGACGATG	260
Qy	1099	TTAAATAAACTTCAACAATGCACATACCGGACACCGCCCATCTGAGACCACTGCTGAG	1158
Db	261	GATGATGATGACGACGATGATGATGATGATGACGACGACGATGATGACGAGTACGAC	320
Qy	1159	GGCAGGGCGAGACGGACAGTGCATCTCAGCTTCAAAATGATGACAAATGTTTCGATGAC	1218
Db	321	GATGATGATGATGACAAAGAAAGATGATGATGACGACGATGACGACGATGATGATGACGAC	380

RESULT 15	
CD099026	752 bp mRNA linear EST 15-MAY-2003
LOCUS	
DEFINITION	AGENCOURT_14011375 NICHD_XGC_Tad1 Xenopus laevis cDNA clone
	IMAGE:6939285 5', mRNA sequence.
ACCESSION	CD099026
VERSION	CD099026.1 GI:30752129
KEYWORDS	EST.
SOURCE	Xenopus laevis (African clawed frog)
ORGANISM	Xenopus laevis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE	1 (bases 1 to 752)
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs- <a href="mailto:email.nih.gov">email.nih.gov</a> Tissue Procurement: Drs. Donald Brown and Liguang Cai cDNA Library Preparation: CLONTECH cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LNCM3263 row: n column: 20 High quality sequence stop: 496.

**FEATURES**  
**SOURCE**

## ORIGIN

Query Match	2.9%	Score 62.8;	DB 14;	Length 752;
Best Local Similarity	56.9%;	Pred. No. 74;		
Matches 115;	Conservative	0;	Mismatches 87;	Indels 0;
				Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:21:32 ; Search time 57 Seconds  
(without alignments)  
2027.402 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTAEFLKPFITPTSGNDGF.....DRNAVLREGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2162	100.0	409	5	AAU77718	Aau77718 Drosophil
2	2082.5	96.3	406	5	AAU77717	Aau77717 Drosophil
3	2045	94.6	409	5	AAU77716	Aau77716 Drosophil
4	1656	76.6	325	4	ABB67553	Abb67553 Drosophil
5	149	6.9	27	5	AAU88008	Aau88008 Tumour ne
6	149	6.9	27	5	AAU88018	Aau88018 Tumour ne
7	141.5	6.5	330	5	ABG32871	Abg32871 Mouse Zis
8	141	6.5	428	5	ABG93245	Abg93245 C. albica
9	138.5	6.4	1192	3	AA18165	Aab18165 Plasmodiu
10	138.5	6.4	1883	4	ABB60038	Abb60038 Drosophil
11	137	6.3	330	3	AA170227	Aay70227 Human RNA
12	137	6.3	532	4	ABB67173	Abb67173 Drosophil
13	137	6.3	2175	4	ABB65698	Abb65698 Drosophil
14	135	6.2	330	4	AAU21828	Aau21828 Novel hum
15	135	6.2	330	4	AAU86946	Aau86946 Human DNA
16	135	6.2	330	7	ADC46469	Adc46469 Human neo
17	134	6.2	466	6	ABR52760	Abf52760 Protein s
18	133.5	6.2	3583	4	ABB64814	Abb64814 Drosophil
19	133	6.2	234	4	AA1894508	Aab94508 Human pro
20	133	6.2	2500	3	AA1818272	Aab18272 Plasmodiu
21	132	6.1	391	5	ABP53104	Abp53104 Human ect
22	132	6.1	391	5	ABB09082	Abb09082 Human EDA
23	132	6.1	391	6	ABR42324	Abf42324 Human EDA
24	132	6.1	391	7	ABO01928	Abf01928 Human EDA
25	132	6.1	391	7	ADC35224	Adc35224 Human TNF

ALIGNMENTS

RESULT 1

AAU77718  
ID AAU77718 standard; protein; 409 AA.

AC AAU77718;

DT 05-JUN-2002 (first entry)

DE Drosophila melanogaster tumour necrosis factor variant 2 (TNFv2).

KW Tumour necrosis factor; TNF; pesticide; crop-damaging insect;  
KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;  
KW sparse hair; sweat gland aberration; endotoxic shock; inflammation;  
KW haemorrhagic necrosis of tumour; cytotoxicity; TNFv2;  
KW obesity-linked insulin resistance.

XX Drosophila melanogaster.

XX Key Location/Qualifiers  
FH Peptide 1..52  
FT Protein 53..409  
FT /label= Signal\_peptide  
FT /label= mature\_TNF  
FT /note= "Tumour necrosis factor. Specifically claimed in claim 36"

US2002012968-A1.

31-JAN-2002.

20-MAR-2001; 2001US-00813329.

21-MAR-2000; 2000US-0190816P.

(CARR/) CARROLL P M.

(CHEN/) CHEN J.

(RAMA/) RAMANATHAN C S.

(XIAO/) XIAO H.

(GUAN/) GUAN B.

(BOWE/) BOWEN M A.

Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;

WPI; 2002-195121/25.

N-PSDB; ABK11680.

XX





Db 121 DGGDGLSDIADDEDDVSYSSVDDYGDYDYMNLKLNNAHTGTTPTSETTAEGEGT 180  
Qy 181 DSASSASNDNVFDDFTSYNAHKKQERKSRSIADVRNEEQNIQNHTELOEKSSNEATS 240  
Db 181 DSASSASNDNVFDDFTSYNAHKKQERKSRSIADVRNEEQNIQNHTELOEKSSNEATS 240  
Qy 241 KESPAPLHRRMRHRSRHLVRK-----ARSEDSRPAAHFHLSSRRRHQSGMYHGDM 294  
Db 241 KE-----RMSHRHRLVRKGESLSARSSEDSRPAAHFHLSSRRRHQSGMYHGDM 291  
Qy 295 YIGNDNERNVSQGHFQTRDGLTVTNTGLYYVYQAICYNNSHDQNGFVFGDTPFLQCL 354  
Db 292 YIGNDNERNVSQGHFQTRDGLTVTNTGLYYVYQAICYNNSHDQNGFVFGDTPFLQCL 351  
Qy 355 NTVPTNMPKHVHTCHTSLIHLERNERHLKDINHNRNAVLRREGNRSYFGIFKV 409  
Db 352 NTVPTNMPKHVHTCHTSLIHLERNERHLKDINHNRNAVLRREGNRSYFGIFKV 406

RESULT 3  
AAU77716  
ID AAU77716 standard; protein; 409 AA.  
XX  
AC AAU77716;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE Drosophila melanogaster tumour necrosis factor (TNF).  
XX  
KW Tumour necrosis factor; TNF; pesticide; crop-damaging insect;  
KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;  
KW sparse hair; sweat gland aberration; endotoxic shock; inflammation;  
KW haemorrhagic necrosis of tumour; cytotoxicity;  
KW obesity-linked insulin resistance.  
XX  
OS Drosophila melanogaster.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..52  
FT Protein /label= Signal\_peptide  
FT /label= mature TNF  
FT /note= "Tumour necrosis factor. Specifically claimed in  
FT claim 34"

US2002012968-A1.  
31-JAN-2002.  
20-MAR-2001; 2001US-00813329.  
21-MAR-2000; 2000US-0190816P.  
(CARR/) CARROLL P M.  
PA (CHEN/) CHEN J.  
PA (RAMA/) RAMANATHAN C S.  
PA (XIAO/) XIAO H.  
PA (GUAN/) GUAN B.  
PA (BOWE/) BOWEN M A.  
XX  
PI Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;  
XX WPI; 2002-195121/25.  
DR N-PSDB; ABK11678.  
XX  
XX New Drosophila tumor necrosis factor molecule, useful in controlling  
PT agriculturally important pests, e.g. comprises modifying the growth,  
PT feeding or reproduction of crop-damaging insects or insects of farm  
PT animals.  
XX

PS Claim 6; Fig 1A-C; 119pp; English.  
XX  
CC The invention describes an isolated tumour necrosis factor polypeptide  
CC (TNF). The polypeptide and polynucleotide are useful in controlling  
CC agriculturally important pests, particularly by modifying the growth,  
CC feeding and/or reproduction of crop-damaging insects or insects of farm  
CC animals. The polypeptide and polynucleotide are useful for modulating  
CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus  
CC the polypeptide and polynucleotide may be useful for treating,  
CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal  
CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-  
CC like disorders, e.g. sparse hair, abnormal or missing teeth or sweat  
CC gland aberrations in animals (e.g. insects and potentially humans),  
CC endotoxic shock, inflammation, haemorrhagic necrosis of tumours,  
CC cytotoxicity and obesity-linked insulin resistance, all of which involve  
CC TNF molecules. This is the amino acid sequence of the Drosophila  
CC melanogaster tumour necrosis factor protein, described in the invention  
XX  
SQ Sequence 409 AA;  
Query Match 94.6%; Score 2045; DB 5; Length 409;  
Best Local Similarity 96.3%; Pred. No. 1.5e-171;  
Matches 394; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
Qy 1 MTAETLKPFTTPTTSANDDGFPAKATSTATAQRTRQLPLVLGFIGLVVAILALTIWQ 60  
Db 1 MTAETLKPFTTPTTSANDDGFPAKATSTATAQRTRQLPLVLGFIGLVVAILALTIWQ 60  
Qy 61 TTRVSHLDKELKSLKRVVDNLQORLGINYLDEFDFQKEYENALIDYPPKVDGLTDEDD 120  
Db 61 TTRVSHLDKELKSLKRVVDNLQORLGINYLDEFDFQKEYENALIDYPPKVDGLTDEDD 120  
Qy 121 DGGDGLSDIADDEDDVSYSSVDDVGADYEDYDMLNKLNAHTGTTPTSETTAEGEGT 180  
Db 121 DGGDGLSDIADDEDDVSYSSVDDVGADYEDYDMLNKLNAHTGTTPTSETTAEGEGT 180  
Qy 181 DSASSASNDNVFDDFTSYNAHKKQERKSRSIADVRNEEQNIQNHTELOEKSSNEATS 240  
Db 181 DSASSASNDNVFDDFTSDALKKQERKSRSIADVRNEEQNIQNHTELOEKSSNEATS 240  
Qy 241 KESPAPLHRRMRHRSRHLVRKARSEDSRPAAHFHLSSRRRHQSGMYHGDMYICNDN 300  
Db 241 KESPAALHLRRMRHRSRHLVRKARSEDSRPAAHFHLSSRRRHQSGMYHGDMYIENDR 300  
Qy 301 ERNSYQGHFOTRGVLTVTNTGLYYVYQAICYNNSHDQNGFIVFGDTPFLQCLNTVPTN 360  
Db 301 ERCSYQGHFOTRGVLTVTNTAGLYYYVYQAIWGYNSHDQNGFIVFGDTPFLQCLNTVPTN 360  
Qy 361 MPKHVHTCHTSGLIHLERNERHLKDINHNRNAVLRREGNRSYFGIFKV 409  
Db 361 MPKHVHTCHTSGLIHLERNERHLKDINHNRNAVLRREGNRSYFGIFKV 409  
RESULT 4  
ABB67553  
ID ABB67553 standard; protein; 325 AA.  
XX  
AC ABB67553;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 29451.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.









KW actinic keratosis; bursitis; arteriosclerosis; artherosclerosis;  
 KW hepatitis; myelofibrosis; primary thrombocythemia; psoriasis; cancer;  
 KW mixed connective tissue disease; MCTD; HIV; uveitis; Crohn's disease;  
 XX allergy; rheumatoid arthritis; parasitic infection.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 2 /note= "Potential phosphorylation site"

FT Domain 9. .40

FT Modified-site 9 /label= Zinc\_finger

FT Modified-site 30 /note= "Potential phosphorylation site"

FT Modified-site 48. .55 /note= "Potential phosphorylation site"

FT Binding-site /label= ATP/GTP Binding\_site\_motif\_A

FT Modified-site 55 /note= "P-loop"

FT Modified-site 65. .94 /note= "Potential phosphorylation site"

FT Domain /label= Zinc\_finger

FT Modified-site 65 /note= "Potential phosphorylation site"

FT Modified-site 90 /note= "Potential phosphorylation site"

FT Modified-site 102 /note= "Potential phosphorylation site"

FT Modified-site 120 /note= "Potential phosphorylation site"

FT Modified-site 124 /note= "Potential phosphorylation site"

FT Modified-site 153 /note= "Potential phosphorylation site"

FT Modified-site 188 /note= "Potential phosphorylation site"

FT Modified-site 193 /note= "Potential phosphorylation site"

FT Modified-site 198 /note= "Potential phosphorylation site"

FT Modified-site 198 /note= "Potential phosphorylation site"

FT Modified-site 214 /note= "Potential phosphorylation site"

FT Modified-site 222 /note= "Potential phosphorylation site"

FT Modified-site 244 /note= "Potential phosphorylation site"

FT Modified-site 244 /note= "Potential phosphorylation site"

FT Modified-site 259 /note= "Potential phosphorylation site"

FT Modified-site 262 /note= "Potential phosphorylation site"

FT Modified-site 268 /note= "Potential phosphorylation site"

FT Modified-site 273 /note= "Potential phosphorylation site"

FT Modified-site 279 /note= "Potential phosphorylation site"

FT Modified-site 288 /note= "Potential phosphorylation site"

FT Modified-site 294 /note= "Potential phosphorylation site"

FT Modified-site 303 /note= "Potential phosphorylation site"

FT Modified-site 326 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

FT Modified-site 327 /note= "Potential phosphorylation site"

XX PD

XX 02-MAR-2000.

XX PF 20-AUG-1999; 99WO-US019361.

XX PR 21-AUG-1998; 98US-0097550P.

XX PR 12-JAN-1999; 99US-0115639P.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Hillman JL, Yue H, Tang YT, Corley NC, Guegler KJ, Gorgone GA;

XX PI Patterson C, Baughn MR, Lai P, Bandman O, Reddy R, Azimzai Y;

XX PI Shih LL, Yang J, Lu DAM;

XX DR WPI; 2000-237651/20.

XX DR N-PSDB; AAZ51257.

XX PT Human RNA-associated proteins useful in diagnosing, treating and

XX PT preventing cell proliferative, autoimmune, inflammatory and infectious

XX PT disorders.

XX PS Claim 1; Page 85-86; 123pp; English.

XX CC The present amino acid sequence is the human RNA-associated protein-8

XX CC (RNAAP-8), identified in Incyte clone 869138, derived from LUNGAST01

XX CC library. It is expressed in reproductive, nervous, gastrointestinal and

XX CC haematopoietic/immune tissues. It has cytoskeletal, immunosuppressive,

XX CC antiinflammatory, antiarteriosclerotic, hepatotropic, keratolytic,

XX CC neuroprotective, antipsoriatic, anti-HIV, antiallergic, antirheumatic,

XX CC virucide, antiarthritic, ophthalmological and antimicrobial activity.

XX CC RNAAP antibodies are useful for diagnosis of diseases associated with

XX CC altered expression or activity of RNAAP. It is used to treat cell

XX CC proliferative, autoimmune, inflammatory and infectious disorders, like

XX CC actinic keratosis, bursitis, arteriosclerosis, atherosclerosis,

XX CC cirrhosis, hepatitis, myelofibrosis, mixed connective tissue disease

XX CC (MCTD), psoriasis, primary thrombocythemia and cancer, HIV, allergies,

XX CC rheumatoid arthritis, uveitis, Crohn's disease, and bacterial, viral and

XX CC parasitic infections

XX SQ Sequence 330 AA;

Query Match 6.3%; Score 137; DB 3; Length 330;

Best Local Similarity 21.2%; Pred. No. 0.0023;

Matches 68; Conservative 49; Mismatches 107; Indels 96; Gaps 12;

QY 20 FPAKATSTATAQRTRQLIPLVLGFIGLGVVA-----ILATTIQTTRVSHLDKELKSL 74

Db 25 FARRTSCNRCGREKTEAKMKAGGTGIGTLAEKSRGLFSANDWQCKTCSNVNWARRE 84

QY 75 KRVD-----NLQORLG-----INLID-----EFDFQKEYENALIDYKVDG 113

Db 85 CNMNTPKYAKLEERTGYGGFNRENVEYIERESDGYEDFGKRRK-----YRGKAVG 140

QY 114 ---LTDDEDDDDGDLDSIADDEDDDVSYSSVDVGDYEDYDMLNKLNAHTGTPTS 170

Db 141 PASILKEVEDKESEEE---EDEDEDLSKYLD----- 170

QY 171 ETATGEGETDSASSASNDNDVDFDFTSYNAHKQKRSRSDIADVRNEQNIQNHTEL 230

Db 171 ----EDEDDADLSKYN----LDASEEDSNKKKSNRRSR----- 204

QY 231 QEKSSNEATSKESPAPLHRRMRHSHRLLVVKARS-----EDSRPAAPHLSRRRHQ 285

Db 205 KRRSHSRSSSRSSSPSSSRSSSRSSSRSSSRSSSRSSSRSSSRSSSRSSSRSSSR 264

QY 286 GSMGYHGDYIGNDN--ERN 303

Db 265 GSSSPKRSYSSSSSSSPERN 284

RESULT 12

ABB67173

ID ABB67173 standard; protein; 532 AA.





RESULT 14  
AAU21828  
ID AAU21828 standard; protein; 330 AA.  
XX  
AC AAU21828;  
XX  
DT 06-DEC-2001 (first entry)  
XX  
DE Novel human neoplastic disease associated polypeptide #261.  
XX  
KW Human; neoplastic disease associated polypeptide; cancer;  
KW hyperproliferative disorder; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW neuroprotective; cytostatic; anti inflammatory; vasotropic.  
XX  
OS Homo sapiens.  
XX  
PN WO200155163-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001358.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229345P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:33:39 ; Search time 18 Seconds  
(without alignments)  
1173.057 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTATLKPFTTTSANDGF.....DRNAVLREGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pap:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pap:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pap:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pap:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pap:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	132	6.1	391	4	US-09-342-681C-2
2	129	6.0	1085	1	US-08-431-080-28
3	129	6.0	1085	2	US-08-938-534-26
4	129	6.0	1085	4	US-09-345-294-26
5	122.5	5.7	706	1	US-08-339-152A-29
6	122.5	5.7	706	2	US-08-007-999B-4
7	122.5	5.7	706	2	US-08-689-276A-4
8	122.5	5.7	714	2	US-08-990-114-3
9	122.5	5.7	714	4	US-09-241-333-3
10	121.5	5.6	296	1	US-07-712-476A-1
11	121.5	5.6	314	4	US-09-134-253-1
12	121.5	5.6	314	4	US-09-206-576-2
13	118.5	5.5	282	1	US-07-712-476A-5
14	116	5.4	249	3	US-09-105-343A-2
15	116	5.4	300	6	5340934-6
16	115.5	5.3	289	4	US-09-589-287B-38
17	115.5	5.3	289	4	US-09-588-947A-38
18	115.5	5.3	289	4	US-09-589-286A-38
19	115.5	5.3	381	4	US-09-216-393B-8
20	115	5.3	10182	4	US-09-134-001C-3159
21	114.5	5.3	187	4	US-09-396-937-8
22	114.5	5.3	414	1	US-07-667-276A-4
23	114.5	5.3	1435	2	US-08-568-459A-4
24	114.5	5.3	1435	2	US-08-487-826B-4
25	114.5	5.3	1435	4	US-09-210-288-4
26	114	5.3	878	4	US-09-540-236-3401
27	113	5.2	226	1	US-08-431-080-26

28	113	5.2	226	2	US-08-938-534-26	Sequence 26, Appl
29	113	5.2	226	4	US-09-345-294-26	Sequence 26, Appl
30	112.5	5.2	553	4	US-09-533-029-90	Sequence 90, Appl
31	112.5	5.2	1162	2	US-08-728-323A-2	Sequence 2, Appl
32	112.5	5.2	1162	4	US-09-298-568-2	Sequence 2, Appl
33	112.5	5.2	1162	4	US-09-410-399-2	Sequence 2, Appl
34	112.5	5.2	1850	4	US-09-620-093A-5	Sequence 5, Appl
35	112	5.2	3696	4	US-09-134-001C-5080	Sequence 5080, Ap
36	111.5	5.2	182	4	US-09-396-937-16	Sequence 16, Appl
37	111.5	5.2	266	4	US-09-589-287B-19	Sequence 19, Appl
38	111.5	5.2	266	4	US-09-879-919-24	Sequence 24, Appl
39	111.5	5.2	266	4	US-09-588-947A-19	Sequence 19, Appl
40	111.5	5.2	266	4	US-09-589-286A-19	Sequence 19, Appl
41	111	5.1	278	6	5340934-8	Patent No. 5340934
42	111	5.1	628	3	US-08-656-034-2	Sequence 2, Appl
43	110.5	5.1	487	4	US-09-386-962C-14	Sequence 14, Appl
44	110	5.1	173	4	US-09-396-937-10	Sequence 10, Appl
45	109.5	5.1	913	4	US-08-971-089-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-09-342-681C-2  
; Sequence 2, Application US/09342681C  
; Patent No. 6355782  
; GENERAL INFORMATION:  
; APPLICANT: Zonana et al.  
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins  
; FILE REFERENCE: 52978  
; CURRENT APPLICATION NUMBER: US/09/342,681C  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 60/092,279  
; PRIOR FILING DATE: 1998-07-09  
; PRIOR APPLICATION NUMBER: 60/112,366  
; PRIOR FILING DATE: 1998-12-15  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-342-681C-2

Query Match	6.1%;	Score 132;	DB 4;	Length 391;
Best Local Similarity	28.6%;	Pred. No. 0.00046;		
Matches	44;	Conservative	31;	Mismatches 63; Indels 16; Gaps 8;
Qy	264	KARSEDRPAAHFHLSSRRRHQGS-MGYHGDWYTG---NDNER---NSYQGHFQTRDGVLT	317	
Db	240	KAGTRENQPAV-VHLOG---QGSAIQVKNLDSGVLDNSRITMNPVKVFLHPRSGELE	294	
Qy	318	VTNTGLYVYVYQAQ--ICYNNSHDQNGFTVFGDTPFLQCLNTVPTNMPHKVHTHTSGLIH	375	
Db	295	VLVDGTGYFYISQVEVYINFTDFASYEVWVDEKPFLOCTRSIETGKTN-YNTCVTAGVCL	353	
Qy	376	LERNERHLDIHDNRNAVLRGNRSYFGIFKV	409	
Db	354	LKARQKIAVKNVHADIS--INMSKHITFFGAIRL	385	

RESULT 2  
US-08-431-080-28  
; Sequence 28, Application US/08431080  
; Patent No. 5698686  
; GENERAL INFORMATION:  
; APPLICANT: Gottschling, Daniel E.  
; APPLICANT: Singer, Miriam S.  
; TITLE OF INVENTION: Telomerase Compositions and Methods  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee



RESULT 4  
US-09-345-294-28  
; Sequence 28, Application US/09345294  
; Patent No. 6387619  
; GENERAL INFORMATION:  
; APPLICANT: Gottechtling, Daniel E.  
; SINGER, Miriam S.  
; TITLE OF INVENTION: Telomerase Compositions and Methods  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: TEXAS  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/345,294  
; FILING DATE: 30-Jun-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/431,080  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, David L.  
; REGISTRATION NUMBER: 32,165  
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (713) 789-2679  
; TELEX: 79-0924  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1085 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-345-294-28

Query Match 6.0%; Score 129; DB 4; Length 1085;  
Best Local Similarity 19.9%; Pred. No. 0.0039;  
Matches 73; Conservative 48; Mismatches 126; Indels 120; Gaps 12;  
Qy 69 KELSLKRVVNLQRLGYNLDFDFEQ-----KEYENALIDYPKKVDGLTDEDDDDG 123  
Db 94 KSKLIRQIUNDEG-----TSSDQAVTDGESENEEESEEDDEDDDDDD 147  
Qy 124 DGLDSIADDEDDVSYSDVDGADYED-----YDMLNKLN 160  
Db 148 DGSDDSDSE-----TSSDDENIDFVKLTARKKRAMKALSAMNTNTLYSSKNSNK 201  
Qy 161 NAHTGTPTSTTAE-----GEGETDSASSANDNVFDDF 196  
Db 202 NKSVKLSPKNEEEFEQKEEKEEKEEQKQESNKKYVNGSGTTTQOALSFKFKEDDG 261  
Qy 197 TSY-NAHKKKQERKSRISADVRNEQN-----IQNHYTELOEKSNEATSKESP 244  
Db 262 ISFGNGEGYNEIDGEEVDLKNKENNGNEEDKLDKVMGLNDELPPNISEDESE-- 319  
Qy 245 APLHRRMRHRRHLLVRKARSDSRPAAPHFLSSRRRHQSGMGY-----HGDYI GND 299  
Db 320 -----YDIDQDAVFDVINNEDSHGEIGTDLTGEDDLFILE 356  
Qy 300 NERN-----SYOG--HFQTRDGLVTVTNTGLYYVYAO-ICYNNSHDQNGFIVFQ 345  
Db 357 EEQNI VSELQNDLSFDGSIHEEGSDFEVDAENKFLQNEYNQENGVDEEDDEDEIMSD 416

Qy 346 GDTFFFLQ 352  
Db 417 FDMPEYE 423  
RESULT 5  
US-08-339-152A-29  
; Sequence 29, Application US/08339152A  
; Patent No. 5643726  
; GENERAL INFORMATION:  
; APPLICANT: Tanzi, Rudolph E.  
; APPLICANT: Kovacs, Dora M.  
; TITLE OF INVENTION: Methods For Modulating Transcription  
; NUMBER OF SEQUENCES: 33  
; TITLE OF INVENTION: From The Amyloid -Protein Precursor (APP) Promoter  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Ave., NW, Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/339,152A  
; FILING DATE: 10-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steffe, Eric K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 0609.4120000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 706 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-339-152A-29

Query Match 5.7%; Score 122.5; DB 1; Length 706;  
Best Local Similarity 21.1%; Pred. No. 0.0084;  
Matches 63; Conservative 36; Mismatches 105; Indels 95; Gaps 13;  
Qy 91 DEFDFQKEYENALIDY-----PKKVD-----GLTDEEDDDGGLDSIADDEDDV 137  
Db 218 EEEDEEESEEDDYVYKSEFPTEADLEDTAAVDEDEDEEEGVV---EORDY 274  
Qy 138 SYSSVDDVGADYEDYTDMLNKLNNNAHTGTPTTSETTAEGE-----GETDSASSAND 189  
Db 275 YDTEFK--GDYNEE---NPTPEGSDGTMSDKXETHDVKVPPPTPLPTNDVYVFETSD 328  
Qy 190 DNVFDDFTS-----YNAHKKKQER-----KSR 212  
Db 329 DNEHARFQKAKEQLIERHNRMDRVKKWEAEALQAKLPKAEQRTLIQHFOAMVKALE 388  
Qy 213 IADVRNEEQIQNHYTELOEKSNE-----ATSKESPALPHHR----- 250  
Db 389 KAAASQQLVETHLARVEMLMNRRMALENYLAALQSRDPPRPHILQPLRYVRAEN 448  
Qy 251 -RRMHS--RH--RHLIV-----RKARSEDSPAAHFLSSRRRHQSGMGYHGMYIGNDNERN 303  
Db 449 KRLHTIRHYQHVLAVDPEKAAQMSQVMTHLHVBERRNQSLSLLYKDPVVARIQEND 507  
RESULT 6  
US-08-007-999B-4



; Sequence 4, Application US/08007999B  
; Patent No. 5851787  
; GENERAL INFORMATION:  
; APPLICANT: Wasco, Wilma  
; APPLICANT: Bupp, Keith  
; APPLICANT: Magendantz, Margaret  
; APPLICANT: Tanzi, Rudolph  
; APPLICANT: Solomon, Frank  
; TITLE OF INVENTION: AMYLOID PRECURSOR-LIKE PROTEIN AND USES  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.  
; STREET: 1100 New York Ave., NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/007,999B  
; FILING DATE: 21-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/872,642  
; FILING DATE: 20-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/930,022  
; FILING DATE: 17-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, G. Kevin  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: 0609.3520002/JAG/GKT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)371-2571  
; TELEFAX: (202)371-2540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 706 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-007-999B-4

Query Match 5.7%; Score 122.5; DB 2; Length 706;  
Best Local Similarity 21.1%; Pred. No. 0.0084;  
Matches 63; Conservative 36; Mismatches 105; Indels 95; Gaps 13;  
QY 91 DEFDFQKEYENALIDY-----PKKVD-----GLTDEEDDDGGLDSIADDEDDV 137  
DB 218 EEEDEEEEDDEEDYVYKSEPTTEADLEDFTEAAVDEDEDEGEVYV---EDRDY 274  
QY 138 SYSSVDDVGADYEDYDMLNKLNNNAHTGTTPTSETTAEGE-----GETDSASSAND 189  
DB 275 YYDTFK--GDDYNEE-----NPTFPGSDGTMDSKEITHDVKVPPTPLPTNDVDVYFETSAD 328  
QY 190 DNVFDDFTS-----YNAHKKQER-----KSRSS 212  
DB 329 DNEHARFQAEKEQLIEHRNRMDRVKKEWEAELOAKNLPKAEROTILQHFOAMVKALE 388  
QY 213 IADVRNEEQINQNHTELQEKSSNE-----ATSKESPAPLHHR-----250  
DB 389 KAEAASEKQQLVETHLARVEAMLNDRMALENYLAALQSRDPPRPHRLQLPLRRYVRAEN 448  
QY 251 -RRMHS-RH-RHLLV---RKARSDSPAAHFLSSRRHQSNGYHGMVIGNDNERN 303  
DB 449 KDLRLTIRHYOHVLAVDPEKAAQMKSQVMTHLVIEERNQSLSLLYKDPYVARIQEND 507

RESULT 7  
US-08-689-276A-4

; Sequence 4, Application US/08689276A  
; Patent No. 5891991  
; GENERAL INFORMATION:  
; APPLICANT: Wasco, Wilma  
; APPLICANT: Bupp, Keith  
; APPLICANT: Magendantz, Margaret  
; APPLICANT: Tanzi, Rudolph  
; APPLICANT: Solomon, Frank  
; TITLE OF INVENTION: AMYLOID PRECURSOR-LIKE PROTEIN AND USES  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Ave., NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/689,276A  
; FILING DATE: 06-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/007,999  
; FILING DATE: 21-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/872,642  
; FILING DATE: 20-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/930,022  
; FILING DATE: 17-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JORGE A. GOLDSTEIN  
; REGISTRATION NUMBER: 29,021  
; REFERENCE/DOCKET NUMBER: 0609.3520003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)371-2571  
; TELEFAX: (202)371-2540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 706 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-689-276A-4

Query Match 5.7%; Score 122.5; DB 2; Length 706;  
Best Local Similarity 21.1%; Pred. No. 0.0084;  
Matches 63; Conservative 36; Mismatches 105; Indels 95; Gaps 13;  
QY 91 DEFDFQKEYENALIDY-----PKKVD-----GLTDEEDDDGGLDSIADDEDDV 137  
DB 218 EEEDEEEEDDEEDYVYKSEPTTEADLEDFTEAAVDEDEDEGEVYV---EDRDY 274  
QY 138 SYSSVDDVGADYEDYDMLNKLNNNAHTGTTPTSETTAEGE-----GETDSASSAND 189  
DB 275 YYDTFK--GDDYNEE-----NPTFPGSDGTMDSKEITHDVKVPPTPLPTNDVDVYFETSAD 328  
QY 190 DNVFDDFTS-----YNAHKKQER-----KSRSS 212  
DB 329 DNEHARFQAEKEQLIEHRNRMDRVKKEWEAELOAKNLPKAEROTILQHFOAMVKALE 388  
QY 213 IADVRNEEQINQNHTELQEKSSNE-----ATSKESPAPLHHR-----250  
DB 389 KAEAASEKQQLVETHLARVEAMLNDRMALENYLAALQSRDPPRPHRLQLPLRRYVRAEN 448  
QY 251 -RRMHS-RH-RHLLV---RKARSDSPAAHFLSSRRHQSNGYHGMVIGNDNERN 303  
DB 449 KDLRLTIRHYOHVLAVDPEKAAQMKSQVMTHLVIEERNQSLSLLYKDPYVARIQEND 507







Db 212 VSGLLALRPGSSLRIRTL---PWAHLKAAPELTTFGLFOV 248

RESULT 15

5340934-6  
; Patent No. 5340934  
; APPLICANT: TERMINE, JOHN D.; YOUNG, MARIAN F.; FISHER, LARRY W.  
; ROBEY, PAMELA G.  
; TITLE OF INVENTION: CDNA SEQUENCES OF HUMAN BONE MATRIX PROTEINS  
; NUMBER OF SEQUENCES: 13  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/432,044  
; FILING DATE: 03-NOV-1989  
; SEQ ID NO: 6:  
; LENGTH: 300  
5340934-6

Query Match 5.4%; Score 116; DB 6; Length 300;  
Best Local Similarity 23.1%; Pred. No. 0.0099;  
Matches 75; Conservative 32; Mismatches 101; Indels 116; Gaps 17;  
QY 41 VLGFITGLGVVAILALTIWOTRVSHLDKELKSLKRVVDNLQORLGINYLDEF-----D 94  
Db 5 VICFLLGLGTCAI---PVKQADSGSSEKQLYN-----KYPDAVATWLNPD 47  
QY 95 EFQKEYENALIDYPKK-----VDGLTDEEDDDDDGCLDSI---ADDEDD-DVSYSSV 142  
Db 48 PSQKQNLAPQTLPSKSNESHDMDDDEDDDDHVDSDQSDIDSDNDSDVDVDDTDYSHQSD 107  
QY 143 DDVGADYED--YTDMINKL--NNAHTGTTPTSETTAGEGET----- 180  
Db 108 ESHHSDSEDELVTPFDLPATEVFPVPTVD--TYDGRGDSVVYGLRSKSKPRRPDIQ 166  
QY 181 -----DSASSANDNDNVFDDFTSYNAHK 203  
Db 167 YPDATDEITSHMESEELNGAYKAIPVAQDINAPSDWDSDRGKDSYETSQLDD-QSAETHS 225  
QY 204 KQERKRSIADVRNEEQNIQGNHTELQEKSSNEATSKESPAPLHRRRRMHSRHLVV- 262  
Db 226 HKQRLYKRKA--NDESN---EHSVDVIDSQELSKVSRE-----FHSHEFHS-HEDMLVV 273  
QY 263 -RKARSDSRPAAHFHLSSRRHQ 285  
Db 274 DPKSKEEDK-----HLKFRISHE 291

Search completed: September 13, 2004, 10:36:16  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:35:34 ; Search time 131 Seconds  
(without alignment)  
1001.237 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTATLKPFTTPTTSANDGF.....DRNAVLREGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues  
Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Query		Match		Length		DB		ID		Description	
Result No.	Score	Match	Length	DB	ID	DB	ID	DB	ID	DB	ID	DB	ID	DB	ID
1	2162	100.0	409	9	US-09-813-329-6	Sequence 6, Appli									
2	2082.5	96.3	406	9	US-09-813-329-4	Sequence 4, Appli									
3	2045	94.6	409	9	US-09-813-329-2	Sequence 2, Appli									
4	149	6.9	27	9	US-09-813-329-54	Sequence 54, Appli									
5	149	6.9	27	9	US-09-813-329-64	Sequence 64, Appli									
6	145.5	6.7	511	16	US-10-437-963-148275	Sequence 148275,									
7	141	6.5	428	16	US-10-451-467A-448	Sequence 448, App									
8	140.5	6.5	406	16	US-10-437-963-154595	Sequence 154595,									
9	137	6.3	699	16	US-10-408-765A-434	Sequence 434, App									
10	135	6.2	330	9	US-09-764-848-34	Sequence 34, Appli									
11	135	6.2	330	14	US-10-116-016-34	Sequence 34, Appli									
12	135	6.2	330	14	US-10-103-313-555	Sequence 555, App									
13	135	6.2	330	14	US-10-222-020-34	Sequence 34, Appli									
14	132	6.1	391	9	US-09-813-329-8	Sequence 8, Appli									
15	132	6.1	391	10	US-09-729-658B-2	Sequence 2, Appli									

ALIGNMENTS

RESULT 1  
US-09-813-329-6  
; Sequence 6, Application US/09813329  
; Patent No. US20020012968A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Subb Company  
; TITLE OF INVENTION: No. US20020012968A1e1 Drosophila Tumor Necrosis Factor Class Mole  
; TITLE OF INVENTION: Variants Thereof  
; FILE REFERENCE: D0016.np  
; CURRENT APPLICATION NUMBER: US/09/813,329  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,816  
; PRIOR FILING DATE: 2000-03-21  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-813-329-6

Query Match				100.0%; Score 2162; DB 9; Length 409;			
Best Local Similarity				100.0%; Pred. No. 4.1e-167;			
Matches 409; Conservative				0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MTAETLKPFTTPTTSANDGFPAKATATATAQRTRQLPLVLGFIGLGVAILALTIWQ	60				
Db	1	MTAETLKPFTTPTTSANDGFPAKATATATAQRTRQLPLVLGFIGLGVAILALTIWQ	60				
Qy	61	TTTRVSHLDKELKSLKRVDNLQORLNGYLDEFDFQKEYENALIDYPPKVDGLTDEDD	120				
Db	61	TTTRVSHLDKELKSLKRVDNLQORLNGYLDEFDFQKEYENALIDYPPKVDGLTDEDD	120				
Qy	121	DGGLDLSIADDEDDVSYSSVDVGVADYEDYDMLNKLNAHTGTTPTSETTAEGGET	180				
Db	121	DGGLDLSIADDEDDVSYSSVDVGVADYEDYDMLNKLNAHTGTTPTSETTAEGGET	180				
Qy	181	DSASSASNDNDVDFDEFTSYNAHKKQERKRSIADVRNEPQNIQGNHTELQEKSSNEATS	240				



QY	108	PKVDGLTDEDDDDGGGLDSIAADDDDDVSYSSVDVGADYEDYDTMLNKLNNATGTT	167
Db	173	PKKV-ASSDKGEEDSKGSESDSGESDLSLSSKSD-----DTRRK-KKGRKGSH	221
QY	168	PTSET-----TAEGGETDSASSASNDNDNVFDDFTSYNAHKKKOERKSRSIADVNE	219
Db	222	RSSKRSRHRHRHSHSDTEGD-DNSKAEEDSEGSVDSDSMDRRKKRSR-----RHK	272
QY	220	EQNIQGNHTELQEKSSNE-----ATSKESPAPLHRRRHMHSRHRLLVVKAR	266
Db	273	KSKRRGRSRRKKRKNDSNTASEGSSEBEAAVAAGSGSPSLRDSKKKRSRRKRKQSD	332
QY	267	SEDSRPA 273	
Db	333	SEDQAPS 339	
RESULT 7			
US-10-451-467A-448			
; Sequence 448, Application US/10451467A			
; Publication No. US20040161840A1			
; GENERAL INFORMATION:			
; APPLICANT: CONTRERAS, ROLAND HENRI			
; APPLICANT: EBERHARDT, INES			
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS			
; APPLICANT: REEKMAN, RIEKA JOSEPHINA			
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN			
; TITLE OF INVENTION: YEAST AND FUNGI			
; FILE REFERENCE: JAB-1667			
; CURRENT APPLICATION NUMBER: US/10/451.467A			
; CURRENT FILING DATE: 2003-06-19			
; PRIOR APPLICATION NUMBER: EP 00870318.3			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: EP 018700002.1			
; PRIOR FILING DATE: 2001-01-04			
; PRIOR APPLICATION NUMBER: EP 018700003.9			
; PRIOR FILING DATE: 2001-01-09			
; NUMBER OF SEQ ID NOS: 732			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 448			
; LENGTH: 428			
; TYPE: PRT			
; ORGANISM: Candida albicans			
US-10-451-467A-448			
Query Match 6.5%; Score 141; DB 16; Length 428;			
Best Local Similarity 20.8%; Pred. No. 0.0063;			
Matches 76; Conservative 68; Mismatches 138; Indels 84; Gaps 16;			
QY	36	QLIPLV-LGFIGLGVVAILALTIWQTRVYSHLDKEL-----KSLKRVVDNL	81
Db	66	QLLPITAKMSNTQDLVLAYINDVYSRNEELSKLALSKFLACKELPKVSKQLESIIDEV	125
QY	82	QQR-----LGINYLDFDFQKEYENALIDYPKK---VDGLTDEDDDDGGGLDSIAADDD	134
Db	126	ENQEKSKPRNSSDSDSSSESSTSDSSSDSDSDSSSDSDSSSDSDSSSDSDSSDSED	185
QY	135	DDVSYSSVD-DVGADYEDYDTMLN-KUNNAHTGTTPTSETTAGEGETOSAGSASNDNV	192
Db	186	SDDEEDKEDKEAKDNKDSEDSENEKVEEDNKDTSSDSSSDSKSDSDSDSSSDSSSS	245
QY	193	FDDFTSYNAHKKKOERKSRSIADVNE---EQNIQGNHTEL---QEKSSNEATSKESPAP	246
Db	246	DSDDSSSDSDSSSDSDSSSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSSDSE	305
QY	247	LHRRRRMHSR--HRHLLVVKARSE-----DSRPAA-----HFHLSRRR- 283	
Db	306	PEDKKRKHDTDDIKEEKPVKKFKNESESSASSSTSDSIPTAPELPKQPKHFSIDRSKV	365
QY	284	-----HQSGMGVHGD-----YTGNDN--BRNSYQGHFQTRDGLVTV	318
Db	366	NFENSVLQDNTYKGAAGTWGEKASEKLLQVRGKDFTKNKNMKRGYK-----GSITL	419







;  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02

;  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,080  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,414  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,244  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,064

```

; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08

Query Match          6.2%; Score 135; DB 14; Length 330;
Best Local Similarity 21.2%; Pred. No. 0.014;
Matches 68; Conservative 48; Mismatches 108; Indels 96; Gaps 12;

Qy 20 FPAKATATATQRTQLIPLVLGFIGLGLVA-----ILALTWTQTRVSHLDKELKSL 74
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 35 FARTSCNRCGREKTEAKMKAGTGTEIGTKLAEKSRGLFSANDWQCKTCSNVNWARRE 94
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 75 KRVD-----NLQRLG-----INLYD-----EPDEFQKEYENALIDPKKVDG 113
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 95 CNMCNTPKYAKLEERTGYGGFNERNVEYIEREESDGEYDFGKKKK-----YRGKAVG 150
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 114 ---LTDEDDDDGGLDSDIADDDSVSYSDVDGADYEDYDMLNKLNNHAHTGTTPTS 170
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 PASILKEVEDKESGEE---EDEDELSKYXD-----DEDELSKYXD----- 180
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 171 ETTAAGEGETDASASANDNDNVDDFTSYNAHKKKQERKSRIADVRNEEQIQQNHTEL 230
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 ----EDEDEDADLSKN---LDASEREEDSKKKSNRRSR----- 214
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 231 QEKSSNEATSKESAPLHRRMRHSRHLHLVRKARS-----EDSRPAAHFLSSRRRHQ 285
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 KRSRSHSRSSRSRSPSSRSRSRSSRSRSSSSSSQSRSSRSRGRSRSSRSRSHR 274
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 286 GSMGVHGDYMGNDN---ERN 303
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 275 GSSSPRKSSYSSSSSSSPERN 294
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-09-813-329-8
; Sequence 8, Application US/09813329
; Patent No. US20020012968A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Subb Company
; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole
; FILE REFERENCE: D0016 nd
; CURRENT APPLICATION NUMBER: US/09/813,329
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-813-329-8

Query Match          6.1%; Score 132; DB 9; Length 391;
Best Local Similarity 28.6%; Pred. No. 0.03;
Matches 44; Conservative 31; Mismatches 63; Indels 16; Gaps 8;

```

```

Qy 264 KARSEDSRPAAHFLSSRRRHQGS--MGYHGDYMGVIG--NDNER---NSYQGHFQTRDGVLT 317
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 KAGTRENQPAV-VHLQG-----QGSAIQVKNDLSGGVLNDMSRITMNPVKFKLHPRSGELE 294
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 318 VTNTGLYYVYQA--ICYNNSHDQNGFIVFGDTPFLQCLNTVPTNMPHKVHTCHTSLGIH 375
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 VLVDGTYFYISQVEVYINFTDFASYEVVVVDEKPFLLQCTRSIETGKTN-YNTCYTAGVCL 353
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 376 LERNERIHLDIHDRNAVLRREGNNRSYFGIFKV 409
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 LKARQKIAVQVWHDIS--INMSKHTTFFGAIRL 385
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-09-729-658B-2
; Sequence 2, Application US/09729658B
; Publication No. US20030023991A1
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729,658B
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-658B-2

Query Match          6.1%; Score 132; DB 10; Length 391;
Best Local Similarity 28.6%; Pred. No. 0.03;
Matches 44; Conservative 31; Mismatches 63; Indels 16; Gaps 8;

Qy 264 KARSEDSRPAAHFLSSRRRHQGS--MGYHGDYMGVIG--NDNER---NSYQGHFQTRDGVLT 317
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 KAGTRENQPAV-VHLQG-----QGSAIQVKNDLSGGVLNDMSRITMNPVKFKLHPRSGELE 294
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 318 VTNTGLYYVYQA--ICYNNSHDQNGFIVFGDTPFLQCLNTVPTNMPHKVHTCHTSLGIH 375
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 VLVDGTYFYISQVEVYINFTDFASYEVVVVDEKPFLLQCTRSIETGKTN-YNTCYTAGVCL 353
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 376 LERNERIHLDIHDRNAVLRREGNNRSYFGIFKV 409
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 LKARQKIAVQVWHDIS--INMSKHTTFFGAIRL 385
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: September 13, 2004, 10:46:53
Job time : 133 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:32:13 ; Search time 19 Seconds  
(without alignments)  
2070.649 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTATLKPFTTPTSANDGF.....DRNAVLRGNRRSYGIFKV 409

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165.5	7.7	3268	2 S69625	hypothetical prote
2	153.5	7.1	852	2 A34373	histidine-rich cal
3	144.5	6.7	764	2 S64951	hypothetical prote
4	138.5	6.4	1192	2 A71623	probable secreted
5	138.5	6.4	1883	2 T13944	chromodomain helic
6	138.5	6.4	3097	2 T28635	glutamate synthase
7	137	6.3	699	2 A54660	histidine rich cal
8	137	6.3	2175	1 S03170	homeotic protein c
9	135	6.2	662	2 T18233	probable transcrip
10	134	6.2	466	2 S19365	hypothetical prote
11	133	6.2	2500	2 G71609	hypothetical prote
12	130.5	6.0	775	2 T03777	hypothetical prote
13	129	6.0	692	2 T03777	homeotic protein H
14	129	6.0	1085	2 S55352	IFH1 protein - yea
15	127	5.9	872	2 T18861	probable peptide-a
16	126.5	5.9	1274	2 A89959	hypothetical prote
17	126	5.8	463	2 T41390	zinc finger protei
18	125	5.8	432	2 G71621	MAK16 homolog PFB0
19	125	5.8	709	2 A96349	F8K7.2 protein - A
20	125	5.8	1672	2 T46237	hypothetical prote
21	125	5.8	2910	2 T28156	DNA-directed RNA p
22	124.5	5.8	3844	2 T18402	asparagine/asparta
23	123.5	5.7	487	2 T10215	hypothetical prote
24	123.5	5.7	630	2 H96770	protein heat shock
25	123.5	5.7	1390	2 T31353	polyprotein - Arab
26	122.5	5.7	406	2 H84590	hypothetical prote
27	122.5	5.7	713	2 A27441	nucleolin - Chines
28	122.5	5.7	1001	2 T16419	hypothetical prote
29	122.5	5.7	1379	2 S64603	YTA7 protein - yea

30	121.5	5.6	314	1 S09575	osteopontin precu
31	121.5	5.6	699	2 G86311	hypothetical prote
32	121	5.6	292	2 T30321	hypothetical prote
33	120.5	5.6	675	2 T03744	myoD protein inhib
34	120.5	5.6	964	2 S06028	gene suppressor-of
35	120.5	5.6	5138	2 B96695	hypothetical prote
36	120	5.6	294	1 A37818	osteopontin precu
37	120	5.6	6713	2 B89921	osteopontin precu
38	119.5	5.5	303	1 GEPGO	osteopontin precu
39	119.5	5.5	1078	2 T18352	protein p120 - Myc
40	119	5.5	665	2 B71609	hypothetical prote
41	119	5.5	770	2 H84463	hypothetical prote
42	118.5	5.5	1046	2 T29776	hypothetical prote
43	118.5	5.5	1105	2 T18295	Ap-3 adaptor compl
44	118	5.5	1067	2 S35423	protein kinase egg
45	117.5	5.4	290	2 G86155	hypothetical prote

ALIGNMENTS

RESULT 1

S69625  
hypothetical protein YDR457w - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 23-Mar-2001  
C:Accession: S69625  
R:Dietrich, F.S.  
submitted to the EMBL Data Library, August 1995  
A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.  
A:Reference number: S69554  
A:Accession: S69625  
A:Molecule type: DNA  
A:Residues: 1-3268 <DIE>  
A:Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64910.1; PID:g927738; MIPS:YDR457w  
C:Genetics:  
A:Gene: SGD:TOM1  
A:Cross-references: SGD:S0002865; MIPS:YDR457w  
A:Map position: 4R

Query Match	7.7%	Score 165.5;	DB 2;	Length 3268;
Best Local Similarity	19.4%;	Pred. No. 0.028;		
Matches	78;	Conservative	66;	Mismatches 142; Indels 117; Gaps 16;
Qy	51	VAILALTIWTVSHLDKELSKRVVDNLQORLGI	----	NYLDEFDEFOKE-----YE 101
Db	1856	VAVEALNTISSTR	-----	NNFSEHFKIEDHDEVEDESDKEIPDMFK 1900
Qy	102	NALIDYPKKVDGLTD	----	EEDDDD-----GLDSIADDEDDVDVSYS 141
Db	1901	NSAL	-----	GMVDVEDIEEDDDDDTSLIGDDDAFVSDNGFEVVFSDDEDDMGEEED 1953
Qy	142	VDDVGADYEDYDMLKNNAHTGTTPTSTTAAEGEETDSASASANDNVDFD	----	TS 198
Db	1954	ADDAARSDSE-NELSEMSQSTADGTDVDYEVDDADGLIINIDQPSGDDEADYDANIS	----	2012
Qy	199	YNAHKKKQERKSIADVNEEQINQNHTELOE	-----	KSSNEATSKES 243
Db	2013	HSSHSENEDDASMDVIEVYDDELS-SGYDVLSDYDVDESGLSLSDSDESSSE	----	2071
Qy	244	PAPLHRRRRMHSRRHLLVRKAR-SEDSRPAAPHLLSSRRRRHQSGMGYHGDMYIGNDNER	----	302
Db	2072	DEPINSTRMGDSRRRLIAEGVELTDDSQ	-----	GESEED 2106
Qy	303	NSVQGHQTDRDGLVTNTGLYYVYAIQICYNNSHDQNGFIVFQG--DTPFLQCLNTVPTN	----	360
Db	2107	D--RGVFRGIEHIFSNENEPLFRVHDEMRRHRHSINRTHFSAMSAPSLSLLNRGRN	----	2164
Qy	361	MPHKVHTCHTSGLIHLF	-----	RNERIHLKDI 387
Db	2165	QSNLINPLGTGLEQVENDISDQTVAGSGSRPRASHLHFSEV	----	2207



A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: mRNA

A;Residues: 1-1883 <STO>

A;Cross-references: EMBL:L77907; NID:g1448982; PID:g1448983; PIDN:AAC37264.1

C;Genetics:

A;Gene: CHD-1

A;Cross-references: FlyBase:FBgn0016132

C;Superfamily: chromodomain helicase CHD1; chromobox homology

C;Keywords: DNA binding

Query Match 6.4%; Score 138.5; DB 2; Length 1883;  
Best Local Similarity 24.5%; Pred. No. 0.76; Mismatches 77; Indels 45; Gaps 8;  
Matches 52; Conservative 38;

Qy 110 KVDGLTDEDDDDGGGLDSIADDEDDVSYSSVDDVGADYEDYDMLNK-LNNAHTGT- 167

Db 83 KTFGTDQDESSDG-----SSGSDSDSADGPGSDQRNQSNINANTSSSL 128

Qy 168 PTSETTAEGEGETDS-----ASSASND-----NVFDDFTS-----YNAHKKKQE 207

Db 129 PKPEQNEEDNETAGQQQPASDASADESSDSSANVSPTSSSSSEEEEDYRPKTRQA 188

Qy 208 RKSRSIAD-----VRNEEQIQNHTELOEKSNEATS--KSPAPLHRRHRRMHSRHR 258

Db 189 RKPTAAEKSKAPAKPNKTKTWDSDESDSDDEVSTAQKRKPAAATTSSRSLAQOQQ 248

Qy 259 HLLVKARSEDSPAPAHFHLSSRRRHQSGMGV 290

Db 249 RRRVKPFSSDDDDASKRCATRKGAASY 280

RESULT 6  
glutamate synthase (NADH2) (EC 1.4.1.14) - malaria parasite (Plasmodium falciparum)  
C;Species: Plasmodium falciparum  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 03-Jun-2002  
C;Accession: T28635  
R;Cowan, G.M.  
submitted to the EMBL Data Library, April 1998  
A;Reference number: Z20490  
A;Accession: T28635  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-3097 <COW>  
A;Cross-references: EMBL:Y17045; NID:e1286063; PID:e1286064; PIDN:CAA76602.1

A;Gene: GlS

C;Keywords: oxidoreductase

Query Match 6.4%; Score 138.5; DB 2; Length 3097;  
Best Local Similarity 23.7%; Pred. No. 1.4;  
Matches 49; Conservative 33; Mismatches 70; Indels 55; Gaps 10;

Qy 78 VDNLQQRIGINYLDEFBFQKEYENALIDYPKKVDGLTDEDDDDGGGLDSIADDEDDV 137

Db 1622 IDN-DENEDMQIDIDEEDDEYN-----DGSDDSDSDSDSGNDD-SDNDDDDI 1670

Qy 138 -----SYSSVDDVGADYEDYDM-----LNKLNNATGTTPTSETTAEGEGETDS 182

Db 1671 DDLDDIDDLDDVDVD-DLDDVDLDDLDLDDVDVDDVHN-----ADNVNCRKDNHTD- 1724

Qy 183 ASSASNDNVDFDFTSYN-----AHKKK-----QERKSRSIADVNREE 220

Db 1725 -HDYDDNNVXDEFPNNDKSGQNSREKRNKKSNRNRIIYKYQKEGKSYNINKKHNN 1783

Qy 221 QNIQNHTELOEKSNEATSKEAPL 247

Db 1784 NNINKNTKTKKXTHLKKKVKPIPI 1810

RESULT 7  
A54660  
histidine rich calcium binding protein - human

C;Species: Homo sapiens (man)  
C;Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 05-Nov-1999  
C;Accession: A54660  
R;Hofmann, S.L.; Topham, M.; Heieh, C.L.; Francke, U.  
Genomics 9, 656-669, 1991

A;Title: cDNA and genomic cloning of HRC, a human sarcolemmal reticulum protein, and 10  
A;Reference number: A54660; MUID:91244309; PMID:2037293  
A;Accession: A54660

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-699 <HOF>

A;Cross-references: GB:M60052; NID:g183918; PIDN:AAA88071.1; PID:g183919

C;Genetics:

A;Gene: GDB:HRC

A;Cross-references: GDB:126369; OMIM:142705

A;Map position: 19q13.3-19q13.3

C;Keywords: calcium binding

Query Match 6.3%; Score 137; DB 2; Length 699;  
Best Local Similarity 20.7%; Pred. No. 0.29;  
Matches 60; Conservative 38; Mismatches 116; Indels 76; Gaps 10;

Qy 91 DEFDEFQKEYENALIDYPKVD---GLTDEDDDDG------GLDSTADDEDD 135

Db 193 EEEEEEEEEEAETEGHQARRHGRGSEDEVDGHHHGHGSHRHHQHEEDDDDDDD 252

Qy 136 DVSYSSVDDVGADYEDYDMLNKLNNATG-TTPTSETTAEGEGETDSA---SSASNDNN 191

Db 253 DDDDDDDDDVSIYRHOA-----HRHQHGIEEDVDGHHHHRDPSHRHSHEEDDN 305

Qy 192 VFDDFTSYNAHK--KKQERKSRSIADVRNE-----EQNIQGNHTELOEKSNEATSKE 243

Db 306 DDDVSTTEYGHQARRHQDHKEVEAVSGEHHHVPDRHQHGRDEEDEDVSTERWHQG 365

Qy 244 PAPLHH-----RRMHSRHRHLVVRK 264

Db 366 PQVHHGLVDDEEEEBEITVQGHYVASHQPRGHKSDEEDFQDEYKTEVPHHHHVRPRE 425

Qy 265 ARSEDSRPAAHFHLSSRRRHQSGMGVHGDVYGNNDNERNYSY-QGHFOTRD 313

Db 426 EDEVSALGHQAPSHRQSHQDEETHGQR--GSIKEMSHHPGHTTVVKD 473

RESULT 8  
S03170  
homeotic protein cut - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: S03170  
R;Blochliger, K.; Bodmer, R.; Jack, J.; Jan, L.Y.; Jan, Y.N.  
Nature 333, 629-635, 1988

A;Title: Primary structure and expression of a product from cut, a locus involved in spe

A;Reference number: S03170; MUID:88232956; PMID:2897632

A;Accession: S03170

A;Molecule type: mRNA

A;Residues: 1-2175 <BLO>

A;Cross-references: EMBL:X07985; NID:g7767; PIDN:CAA30794.1; PID:g7768

C;Genetics:

A;Gene: cut

A;Cross-references: FlyBase:FBgn0004198

C;Superfamily: homeotic protein cut; cut repeat homology; homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;896-958/Domain: cut repeat homology <CU1>

F;1339-1411/Domain: cut repeat homology <CU2>

F;1617-1689/Domain: cut repeat homology <CU3>

F;1746-1802/Domain: homeobox homology <HOX>

Query Match 6.3%; Score 137; DB 1; Length 2175;  
Best Local Similarity 21.8%; Pred. No. 1.1;  
Matches 64; Conservative 45; Mismatches 111; Indels 74; Gaps 9;

Qy 21 PAKATSTATQRRTRQLPLVLGFIGLGVAVAILATIWQTIVSHLDKELKSLKRVVDN 80









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:23:02 ; Search time 12 Seconds  
(without alignments)  
1774.724 Million cell updates/sec

Title: US-09-813-329-6

Perfect score: 2162

Sequence: 1 MTATLKPFTTSANDGF.....DRNAVLREGNNRSYFGIFKV 409

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	153.5	7.1	852	1 SRCH_RABIT	P16230 oryctolagus
2	137	6.3	699	1 SRCH_HUMAN	P23327 homo sapien
3	137	6.3	2175	1 HMCU_DROME	P10180 drosophila
4	134	6.2	466	1 SRO9_YEAST	P25567 saccharomyc
5	132	6.1	391	1 EDA_BOVIN	Q9beg5 bos taurus
6	132	6.1	391	1 EDA_HUMAN	Q92838 homo sapien
7	129	6.0	1085	1 IFH1_YEAST	P39520 saccharomyc
8	128	5.9	391	1 EDA_MOUSE	O54693 mus musculus
9	123.5	5.7	278	1 OSTP_SHEEP	Q9x8y9 ovis aries
10	123.5	5.7	332	1 Z265_RAT	O35986 rattus norv
11	122.5	5.7	713	1 NUCLE_MESAU	P08199 mesocricetu
12	122.5	5.7	1379	1 TRP7_YEAST	P40340 saccharomyc
13	121.5	5.6	314	1 OSTP_HUMAN	P10451 homo sapien
14	121.5	5.6	326	1 Z265_MOUSE	Q9f020 mus musculus
15	121	5.6	337	1 Z265_HUMAN	O95218 homo sapien
16	120.5	5.6	1042	1 SUWA_DROME	P12297 drosophila
17	120.5	5.6	1161	1 NRDC_MOUSE	O8bhg1 mus musculus
18	120	5.6	294	1 OSTP_MOUSE	P10923 mus musculus
19	119.5	5.5	303	1 OSTP_PIG	P4287 sus scrofa
20	118.5	5.5	1105	1 A3B1_MOUSE	Q9z1t1 mus musculus
21	118.5	5.5	4835	1 MDN1_GIALA	O8c5t1 giardia lam
22	118	5.5	514	1 TUBE_DROVI	Q08171 drosophila
23	118	5.5	1067	1 SGG_DROME	P18431 drosophila
24	117.5	5.4	406	1 SR40_YEAST	P32583 saccharomyc
25	117	5.4	657	1 YGG6_YEAST	P33165 saccharomyc
26	117	5.4	1014	1 HEX2_YEAST	Q00816 saccharomyc
27	116	5.4	249	1 TN12_HUMAN	O43508 homo sapien
28	116	5.4	1025	1 MK21_YEAST	Q12176 saccharomyc
29	116	5.4	4910	1 MDN1_YEAST	Q12019 saccharomyc
30	115.5	5.3	706	1 NUCLE_MOUSE	P09405 mus musculus
31	115.5	5.3	1083	1 HOS4_YEAST	P40480 saccharomyc
32	115	5.3	590	1 YNW7_YEAST	P53863 saccharomyc
33	114.5	5.3	414	1 NSR1_YEAST	P27476 saccharomyc

34	114.5	5.3	441	1 YB66_YEAST	P38193 saccharomyc
35	114.5	5.3	802	1 NAB3_YEAST	P38996 saccharomyc
36	114.5	5.3	1435	1 EBAL_PLAFC	P19214 plasmodium
37	114	5.3	899	1 YMJ3_YEAST	Q04500 saccharomyc
38	113.5	5.2	277	1 OSTK_BOVIN	P31098 bos taurus
39	113.5	5.2	851	1 BCK2_YEAST	P3306 saccharomyc
40	113.5	5.2	1807	1 VTA2_KENLA	P18709 xenopus lae
41	113	5.2	278	1 OSTP_BOVIN	P31096 bos taurus
42	113	5.2	510	1 GAT1_YEAST	P43574 saccharomyc
43	113	5.2	646	1 SGL_BOVIN	P33389 bos taurus
44	112.5	5.2	553	1 WRK6_ARATH	Q9C519 arabidopsis
45	112.5	5.2	1253	1 DSPP_HUMAN	Q9nzw4 homo sapien

## ALIGNMENTS

RESULT 1

ID	SRCH_RABIT	STANDARD;	PRT;	852 AA.
AC	P16230;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Sarcoplasmic reticulum histidine-rich calcium-binding protein precursor.			
DE	HRC OR HCP.			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=Skeletal muscle;			
RX	MEDLINE=90036884; PubMed=2808365;			
RA	Hofmann S.L., Goldstein J.L., Orth K., Moomaw C.R., Slaughter C.A., Brown M.S.;			
RA	"Molecular cloning of a histidine-rich Ca2+-binding protein of sarcoplasmic reticulum that contains highly conserved repeated elements";			
RL	J. Biol. Chem. 264:18083-18090(1989).			
CC	-!- FUNCTION: May play a role in the regulation of calcium sequestration or release in the SR of skeletal and cardiac muscle.			
CC	-!- SUBCELLULAR LOCATION: Sarcoplasmic reticulum lumen.			
CC	-!- SIMILARITY: STRONG, TO HUMAN HRC.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; J05080; AAA31279.1; -.			
DR	PIR; A34373; A34373.			
DR	InterPro; IPR002134; HCP.			
DR	PROSITE; PS00328; HCP; 10.			
KW	Calcium-binding; Signal; Repeat.			
FT	CHAIN 1 27			
FT	SIGNAL 28 852			
FT	POTENTIAL.			
FT	SARCOPLASMIC RETICULUM HISTIDINE-RICH			
FT	CALCIUM-BINDING PROTEIN.			
FT	BLOCKED.			
FT	2 X APPROXIMATE TANDEM REPEATS.			
FT	DOMAIN 59 100			
FT	REPEAT 59 79			
FT	REPEAT 80 100			
FT	DOMAIN 199 470			
FT	REPEAT 199 224			
FT	REPEAT 224 2-1.			
FT	REPEAT 225 2-2.			
FT	REPEAT 254 2-3.			
FT	REPEAT 283 2-4.			
FT	REPEAT 310 2-5.			
FT	REPEAT 311 339			
FT	REPEAT 340 367			



RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Ballew R.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benson P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davidson L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foeller C., Gabriellista A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milashina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).  
 CC -1- FUNCTION: Regulator of cell fate decisions in multiple lineages.  
 CC Specifically, functions as a determination factor that specifies  
 CC sensory organ identity in precursor cells. Probably also involved  
 CC in cell type specification of Malpighian tubules. In absence of  
 CC cut gene external sensory organs are transformed into chordotonal  
 CC organs.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- TISSUE SPECIFICITY: Detected in many cells in the central nervous  
 CC system, all external sensory organs, some peripheral neurons, and  
 CC in the non-neural cells of the spiracles and the Malpighian  
 CC tubules.  
 CC -1- DEVELOPMENTAL STAGE: Cell-specific pattern of expression. Broadly  
 CC expressed during embryonic development.  
 CC -1- DOMAIN: Asn at position 47 of the homeobox may participate in  
 CC regulating DNA-binding activity by promoting homo- and  
 CC heterodimerization.  
 CC -1- SIMILARITY: Belongs to the CUT homeobox family.  
 CC -1- SIMILARITY: Contains 1 homeobox domain.  
 CC -1- SIMILARITY: Contains 3 CUT domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X07985; CAA30794.1; -;  
 CC EMBL; AE003441; AAP46264.2; -;  
 CC PIR; S03170; S03170.  
 CC TRANSFAC; T02004; -;  
 CC FlyBase; FBgn0004198; ct.  
 CC GO; GO:0005634; C:nucleus; IDA.

DR GO; GO:0008597; P:wing margin morphogenesis; NAS.  
 DR InterPro; IPR007108; Cut homeo.  
 DR InterPro; IPR003350; Hmoec CUT.  
 DR InterPro; IPR001356; Homeobox.  
 DR Pfam; PF02376; CUT; 3.  
 DR Pfam; PF00046; homeobox; 1.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX 1; 1.  
 DR PROSITE; PS00071; HOMEBOX 2; 1.  
 DR Transcription regulation; Homeobox; DNA-binding; Coiled coil.  
 KW Developmental protein; Nuclear protein; Repeat; Coiled coil.  
 FT DOMAIN 265 343 COILED COIL (POTENTIAL).  
 FT DOMAIN 433 499 COILED COIL (POTENTIAL).  
 FT DNA BIND 798 964 CUT 1.  
 FT DOMAIN 1056 1161 COILED COIL (POTENTIAL).  
 FT DNA BIND 1329 1417 CUT 2.  
 FT DOMAIN 1463 1522 COILED COIL (POTENTIAL).  
 FT DNA BIND 1608 1695 CUT 3.  
 FT DNA BIND 1745 1804 HOMEBOX.  
 FT DOMAIN 194 210 ALA/GLN-RICH.  
 FT DOMAIN 235 243 ALA-RICH.  
 FT DOMAIN 271 293 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 384 428 ASN-RICH.  
 FT DOMAIN 547 554 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 574 584 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 616 630 ALA-RICH.  
 FT DOMAIN 665 699 HIS/GLN-RICH (OPA-REPEAT).  
 FT DOMAIN 2004 2014 ALA-RICH.  
 FT DOMAIN 2071 2077 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 2124 2136 ALA/PRO-RICH.  
 SQ SEQUENCE 2175 AA; 233628 MW; 08BF80C4861BD0AB CRC64;  
 Query Match 6.3%; Score 137; DB 1; Length 2175;  
 Best Local Similarity 21.8%; Pred. No. 0.6;  
 Matches 64; Conservative 45; Mismatches 11; Indels 74; Gaps 9;  
 Qy 21 PAKATSTATQRRTRQLPVLGVLGVLVAILALTIWQTRVSHLDKELSLKRVVDN 80  
 Db 227 PAATVATGAAAAAATPIATGNVSGSTSNANHT---NSNSHQDEE----- 273  
 Qy 81 LQORLGINYLDEPFOKEVENALIDYPKKVGLTDEEDDDGDLSDIADDEDDVSYS 140  
 Db 274 -----ELDDDEDEDEE-----DEDEENASMQSNADMDLDAQE 311  
 Qy 141 SVDDVGG-----DYEDYD-----MLANKLNNAHGTPTTSETTAEGEGETDS-- 182  
 Db 312 TRTEPSATTQQHQOQDTELEENKQAGEASLVSNHNTTDSNNNSCSRKNNGGNESEQ 371  
 Qy 183 --ASSASNDNVDDFTSYNAHKKQERKSRSIADVRNEQNIGNHTELQEKSSNE--- 237  
 Db 372 HVASSAEDDDCANNTNTSNNTNTSNATSTNT---NNNNNNSSGSEKSKKNNNNNG 429  
 Qy 238 -----ATSKESPAPLHRRHMRHRLLVKARSEDSRPAHFLLSSRRRH 284  
 Db 430 QPAVLLAAKDKETKALLDELQRLRAQEQTHLVQIRLEE-----HLEVKRQH 476  
 RESULT 4  
 SRO9 YEAST STANDARD; PRT; 466 AA.  
 AC P25567;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE SRO9 protein.  
 GN SRO9 OR YCL037C OR YCL37C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OK NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.



```
Best Local Similarity 28.6%; Pred. No. 0.17;
Matches 44; Conservative 31; Mismatches 63; Indels 16; Gaps 8;
QY 264 KARSDSPAAHFLHSSRRRQGS-MGYHGMVIG--NDNER---NSYQGHFQTRDGVLT 317
DB 240 KAGTRENPAV-VHLOG-----QGSNAIQKNDLGGVLDNSGRITWNPVKFKLHPRSGELE 294
QY 318 VTNTGLYVYVYQA--ICYNNSHDQNGFIVFGDTPFLQCLNTVPTNMPHKVHTCHTSGLIH 375
DB 295 VLVDGTGYISQVEVYINFTDFASYEVVVDEKPFLOQTRSIETGKTN-YNTCYTAGVCL 353
QY 376 LERNERHLKDIHNDRAVLREGNRSYFGIFKV 409
DB 354 LKARQKIAVKMVDHDIS--INMSKHTTFFGAIRL 385

RESULT 6
EDA_HUMAN
ID EDA_HUMAN STANDARD; PRT; 391 AA.
AC Q92838; O75910; Q9UP77; Q9Y6L0; Q9Y6L1; Q9Y6L2; Q9Y6L3; Q9Y6L4;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ectodysplasin A (Ectodermal dysplasia protein) (EDA protein).
GN ED1 OR EDA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM I), AND VARIANTS EDA HIS-61 AND LEU-69.
RC TISSUE=Sweat gland;
RX MEDLINE=96331280; PubMed=8696334;
RA Kere J., Srivastava A.K., Montonen O., Zonana J., Thomas N.S.T.,
RA Ferguson B.M., Munoz F., Morgan D., Clarke A., Baybayan P.,
RA Chen E.Y., Ezer S., Saarialho-Kere U., la Chapelle A.,
RA Schlesinger D.;
RT "X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by
RT mutation in a novel transmembrane protein.";
RL Nat. Genet. 13:409-416(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A1), AND VARIANTS EDA.
RC TISSUE=Liver;
RX MEDLINE=98349961; PubMed=9683615;
RA Montreal A.W., Zonana J., Ferguson B.M.;
RT "Identification of a new splice form of the ED1 gene permits
RT detection of nearly all X-linked hypohidrotic ectodermal dysplasia
RT mutations.";
RL Am. J. Hum. Genet. 63:380-389(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS A1; A2; B; C; D; E AND F), AND VARIANTS
RP EDA.
RX MEDLINE=98409495; PubMed=9736768;
RA Bayes M., Hartung A.J., Ezer S., Piepa J., Theeleff I.,
RA Srivastava A.K., Kere J.;
RT "The anhidrotic ectodermal dysplasia gene (EDA) undergoes alternative
RT splicing and encodes ectodysplasin-A with deletion mutations in
RT collagenous repeats.";
RL Hum. Mol. Genet. 7:1661-1669(1998).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS A1 AND C).
RA Clark S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP TISSUE SPECIFICITY, AND ALTERNATIVE SPLICING.
RA Kobielak K., Kobielak A., Trzciak W.H.;
RT "Expression of a novel transcript isoform of the EDA gene in human
RT umbilical cord.";
RL Eur. J. Hum. Genet. Suppl. 7:104-104(1999).
RN [6]
RP RECEPTOR INTERACTION (ISOFORMS A1 AND A2).
RX MEDLINE=20495245; PubMed=11039935;
RA Yan M., Wang L.-C., Hymowitz S.G., Schilbach S., Lee J., Goddard A.,
```

```
RA de Vos A.M., Gao W.-Q., Dixit V.M.;
RT "Two-amino acid molecular switch in an epithelial morphogen that
RT regulates binding to two distinct receptors.";
RL Science 290:523-527(2000).
RN [7]
RP PROCESSING, MUTAGENESIS OF ARG-153, AND CHARACTERIZATION OF VARIANT
RP HIS-156.
RX MEDLINE=21205766; PubMed=11309369;
RA Elomaa O., Pulkkinen K., Hannellius U., Mikkola M., Saarialho-Kere U.,
RA Kere J.;
RT "Ectodysplasin is released by proteolytic shedding and binds to the
RT EDAR protein.";
RL Hum. Mol. Genet. 10:953-962(2001).
RN [8]
RP CHARACTERIZATION OF VARIANTS CYS-155; CYS-156 AND HIS-156, MUTAGENESIS
RP OF ARG-153; LYS-158 AND ARG-159, AND CLEAVAGE SITE.
RX MEDLINE=21309995; PubMed=11416205;
RA Chen Y., Molloy S.S., Thomas L., Gambee J., Baechinger H.P.,
RA Ferguson B.M., Zonana J., Thomas G., Morris N.P.;
RT "Mutations within a furin consensus sequence block proteolytic release
RT of ectodysplasin-A and cause X-linked hypohidrotic ectodermal
RT dysplasia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7218-7223(2001).
RN [9]
RP VARIANT EDA TYR-54.
RX MEDLINE=98292028; PubMed=9630076;
RA Hertz J.M., Noergaard Hansen K., Juncker I., Kjeldsen M.,
RA Gregersen N.;
RT "A novel missense mutation (402C-->T) in exon 1 in the EDA gene in a
RT family with X-linked hypohidrotic ectodermal dysplasia.";
RL Clin. Genet. 53:205-209(1998).
RN [10]
RP VARIANT EDA LYS-63.
RX MEDLINE=98168231; PubMed=9507389;
RA Ferguson B.M., Thomas N.S.T., Munoz F., Morgan D., Clarke A.,
RA Zonana J.;
RT "Scarcity of mutations detected in families with X linked hypohidrotic
RT ectodermal dysplasia: diagnostic implications.";
RL J. Med. Genet. 35:112-115(1998).
RN [11]
RP VARIANT EDA ARG-55.
RX MEDLINE=99399307; PubMed=10469321;
RA Martinez F., Millan J.M., Orellana C., Prieto F.;
RT "X-linked anhidrotic (hypohidrotic) ectodermal dysplasia caused by a
RT novel mutation in ED1 gene: 406T > G (Leu55Arg).";
RL J. Invest. Dermatol. 113:285-286(1999).
RN [12]
RP VARIANTS EDA ARG-60; TYR-252; VAL-269; SER-302 AND MET-378.
RX MEDLINE=21272350; PubMed=11378824;
RA Vincent M.C., Biancalana V., Ginisty D., Mandel J.L., Calvas P.;
RT "Mutational spectrum of the ED1 gene in X-linked hypohidrotic
RT ectodermal dysplasia.";
RL Eur. J. Hum. Genet. 9:355-363(2001).
RN [13]
RP VARIANTS EDA CYS-156; HIS-156; CYS-255; ASP-255; GLY-274; TYR-332 AND
RP THR-349.
RX MEDLINE=21193173; PubMed=11295832;
RA Pasaekkoenen K., Cambiaghi S., Novelli G., Ouzts L.V., Penttinen M.,
RA Kere J., Srivastava A.K.;
RT "The mutation spectrum of the EDA gene in X-linked anhidrotic
RT ectodermal dysplasia.";
RL Hum. Mutat. 17:349-349(2001).
RN CC -1- FUNCTION: Seems to be involved in epithelial-mesenchymal signaling
CC during morphogenesis of ectodermal organs. Isoform A1 binds only
CC to the receptor EDAR, while isoform A2 binds exclusively to the
CC receptor XEDAR.
CC CC
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted.
CC CC
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=8;
CC Comment=Additional isoforms seem to exist;
CC Name=A1; Synonyms=I1;
CC IsoId=Q92838-1; Sequence=Displayed;
```



```

CC CC Name=I;
CC CC IsoId=Q92838-2; Sequence=VSP_006454, VSP_006455;
CC CC Name=A2;
CC CC IsoId=Q92838-3; Sequence=VSP_006464;
CC CC Name=B;
CC CC IsoId=Q92838-4; Sequence=VSP_006462, VSP_006463;
CC CC Name=C;
CC CC IsoId=Q92838-5; Sequence=VSP_006458, VSP_006461;
CC CC Name=D;
CC CC IsoId=Q92838-6; Sequence=VSP_006456, VSP_006457;
CC CC Name=E;
CC CC IsoId=Q92838-7; Sequence=VSP_006459, VSP_006461;
CC CC Name=F;
CC CC IsoId=Q92838-8; Sequence=VSP_006460, VSP_006461;
CC CC ISSUE SPECIFICITY: Not abundant; expressed in specific cell types
CC CC of ectodermal (but not mesodermal) origin of keratinocytes, hair
CC CC follicles, sweat glands. Also in adult heart, liver, muscle,
CC CC pancreas, prostate, fetal liver, uterus, small intestine and
CC CC umbilical chord.
CC CC -!- PTM: N-glycosylated.
CC CC -!- PTM: Processing by furin produces a secreted form.
CC CC -!- DISEASE: Defects in ED1 are the cause of ectodermal dysplasia,
CC CC anhidrotic (EDA) (MIM:305100); also known as X-linked hypohidrotic
CC CC ectodermal dysplasia (XUHED). EDA is a disease characterized by
CC CC sparse hair (atrichosis or hypotrichosis), abnormal or missing
CC CC teeth and the inability to sweat due to the absence of sweat
CC CC glands. EDA is the most common form of over 150 clinically
CC CC distinct ectodermal dysplasias. This disease was already described
CC CC by Darwin.
CC CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC CC -!- SIMILARITY: Contains 1 collagenous domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U59227; AAC50678.1; -
CC CC EMBL; U59228; AAC50679.1; -
CC CC EMBL; AF061189; AAC77371.1; -
CC CC EMBL; AF061190; AAC77372.1; -
CC CC EMBL; AF061191; AAC77373.1; -
CC CC EMBL; AF061192; AAC77374.1; -
CC CC EMBL; AF061193; AAC77375.1; -
CC CC EMBL; AF061194; AAC77376.1; -
CC CC EMBL; AF060999; AAC36302.1; -
CC CC EMBL; AF060998; AAC36303.1; -
CC CC EMBL; AF060992; AAC36303.1; JOINED.
CC CC EMBL; AF060993; AAC36303.1; JOINED.
CC CC EMBL; AF060994; AAC36303.1; JOINED.
CC CC EMBL; AF060995; AAC36303.1; JOINED.
CC CC EMBL; AF060996; AAC36303.1; JOINED.
CC CC EMBL; AF060997; AAC36303.1; JOINED.
CC CC EMBL; AF040628; AAC77363.1; -
CC CC EMBL; AL158069; CAD18890.1; -
CC CC EMBL; AL158141; CAD13493.1; -
CC CC EMBL; HGNC:3157; ED1.
CC CC MIM; 305100; -
CC CC GO; GO:0005856; C:cytoskeleton; TAS.
CC CC GO; GO:0016021; C:integral to membrane; TAS.
CC CC GO; GO:0005624; C:membrane fraction; TAS.
CC CC GO; GO:0005886; C:plasma membrane; TAS.
CC CC GO; GO:0005102; F:receptor binding; TAS.
CC CC GO; GO:0007398; P:ectoderm development; TAS.
CC CC GO; GO:0007165; P:signal transduction; TAS.
CC CC InterPro; IPR008160; Collagen.
CC CC InterPro; IPR006052; TNF family.
CC CC InterPro; IPR008983; TNF_like.
CC CC -----
Query Match 6.1%; Score 132; DB 1; Length 391;
Best Local Similarity 28.6%; Pred. No. 0.17;
Matches 44; Conservative 31; Mismatches 63; Indels 16; Gaps 8;
OY 264 KARSERSRPAHPHLSRRRHQGS-MGYHCDMYIG--NDNER---NSYQGHFOTRGVLT 317
DB 240 KAGTRENQPAV-VHLOG---QGSALQVKNLDSGGVLNDWSRITMNPVKFLHPRSGELE 294
OY 318 VTNTGLYYVYQAQ--ICVNNSHDQNGFIVFOGDTFFLQCLNTVPTNMPHKVHTCHTSLI 375
DB 295 VLVDGTGYFYISQVEVYVYINFTDPAFSEVVDKPFLOCTRSIETGKTN-YNICYTAGVCL 353
OY 376 LERNERHLKDHNDNRNAVLRGNRSYFGIFKV 409
DB 354 LKARQKIAVMQVHADIS--INMSKHTTFFGAIRL 385
RESULT 7
IFH1_YEAST STANDARD; PRT; 1085 AA.
AC P39520;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE IFH1 protein (RRP3 protein)
GN IFH1 OR RRP3 OR YLR223C OR L8083.9.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]_YEAST
RP SEQUENCE FROM N.A.
RC STRAIN=ARCC 28383 / FL100;
RX MEDLINE=95304839; PubMed=7785326;
RA Cheral I., Thuriaux P.;
RT "The IFH1 gene product interacts with a fork head protein in
RL Yeast 11:261-270(1995).
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albertmann K., Andre B., Ansorge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoef A.,
RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.-J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
CC -!- FUNCTION: Controls the pre-rRNA processing machinery in
CC conjunction with FHL1. Could convert FHL1 from a repressor
CC to an activator.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC EMBL; Z29488; CAA82624.1; -
CC CC EMBL; U19027; AAB67412.1; -
CC CC PIR; S55352; S55352.
CC CC GeneOnline; 142285; -.
CC CC SGD; S0004213; IFH1.

```

```
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.
DR GO; GO:0006364; P:rRNA processing; IGI.
KW Nuclear protein; Transcription regulation.
FT DOMAIN 122 163 ASP/GLU-RICH (HIGHLY ACIDIC).
SQ SEQUENCE 1085 AA; 122491 MW; BE1C7DFE06213FE0 CRC64;

Query Match 6.0%; Score 129; DB 1; Length 1085;
Best Local Similarity 19.9%; Pred. NO. 0.85;
Matches 73; Conservative 48; Mismatches 126; Indels 120; Gaps 12;

QY 69 KELSLKRVNQLQRLGINYLDFEDEFQ-----KEYENALIDYPKVGDGLTDEDDDDG 123
Dy 94 KKSLLIQRIQINDDEG-----TESSDYQAVTDGESENEBEESEEDDEDDDDDD 147
QY 124 DGLSDIADDDDDSVSYSDVDGADYED-----YTOMLNLN 160
Dy 148 DGSDDSDSE-----TSDDENIDFVKLTAKRKRAMKALSAMNTNNTLYSRENSNK 201
QY 161 NAHTGTTTPTSETTAB-----GEGETDSASSANDNDNVDFD 196
Dy 202 NKSVKLSPKNEEBEQEKEKEKEEQQESNKKEVSGTITQQAISFKFKKEDDG 261
QY 197 TSY-NAHKKKOERKSRSTADYVNEQN-----IQNHYTELQESNEATSKESP 244
Dy 262 ISFGNGEGYNEDIGEEVLDLNKENNGNEBKLDKSKVMLGNDELRFNISESDESE-- 319
QY 245 APLHRRMRHSHRHLVRKARSDRPAAFHFLSSRRRHQSGMCGY-----HGDWYIGND 299
Dy 320 -----YDIDQDAYFDVINNDSHGEIGTDLTGEDDLPLEE 356
QY 300 NERN-----SYQG--HFQTRDGLVTNTGLVYVYQAQ-ICYNNSHDQNGFVIFQ 345
Dy 357 EEQNIIVSELQNDDELSPDGSITHEGSDPVEDAKENKFLQNEYNGYDEEDDEDEIMSD 416
QY 346 GDTPELQ 352
Dy 417 FDMPPFE 423

RESULT 8
EDA_MOUSE
ID_EDA_MOUSE STANDARD; PRT; 391 AA.
AC 054693; 035705; Q9QJH8; Q9QZ01; Q9QZ02;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ectodysplasin A (EDA protein homolog) (Tabby protein).
GN ED1 OR EDA OR TA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A. (ISOFORMS TAA; TAB AND TAC).
RC STRAIN=129/SV;
RX MEDLINE=98058770; PubMed=9371801;
RA Srivastava A.K., Pispas J., Hartung A.J., Du Y., Ezer S., Jenks T.,
RA Shimada T., Pekkanen M., Mikkola M.L., Ko M.S.H., Thesleff I.,
RA Kere J., Schlesinger D.;
RT "The tabby phenotype is caused by mutation in a mouse homologue of the
RT EDA gene that reveals novel mouse and human exons and encodes a
RT protein (ectodysplasin-A) with collagenous domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:13069-13074(1997).
[2]
RN SEQUENCE FROM N.A. (ISOFORM TAD).
RX MEDLINE=9749184; PubMed=9285798;
RA Ferguson B.M., Brockdorff N., Formstone E., Nguyen T.,
RA Kronmiller J.E., Zonana J.;
RT "Cloning of Tabby, the murine homolog of the human EDA gene: evidence
RT for a membrane-associated protein with a short collagenous domain.";
RL Hum. Mol. Genet. 6:1589-1594(1997).
[3]
RN SEQUENCE FROM N.A. (ISOFORMS TA-A2 AND TA-A3).
```

```
RC TISSUE=Embryo;
RX MEDLINE=20005791; PubMed=10534613;
RA Mikkola M.L., Pispas J., Pekkanen M., Paulin L., Nieminen P., Kere J.,
RA Thesleff I.;
RT "Ectodysplasin, a protein required for epithelial morphogenesis, is a
RT novel TNF homologue and promotes cell-matrix adhesion.";
RL Mech. Dev. 88:133-146(1999).
CC -!- FUNCTION: Involved in epithelial-mesenchymal signaling during
CC morphogenesis of ectodermal organs. Isoform TAA binds only to the
CC receptor EDAR, while isoform TA-A2 binds exclusively to the
CC receptor XEDAR.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Comment=Additional isoforms seem to exist;
CC Name=TAA; Synonyms=A1;
CC IsoId=O54693-1; Sequence=Displayed;
CC Name=TA-A2;
CC IsoId=O54693-2; Sequence=VSP_006471;
CC Name=TA-A3;
CC IsoId=O54693-3; Sequence=VSP_006469, VSP_006471;
CC Name=TAB;
CC IsoId=O54693-4; Sequence=VSP_006466, VSP_006467;
CC Name=TAC;
CC IsoId=O54693-5; Sequence=VSP_006465, VSP_006468;
CC Name=TAD;
CC IsoId=O54693-6; Sequence=VSP_006470;
CC -!- PTM: N-glycosylated (By similarity).
CC -!- PTM: Processing by furin produces a secreted form (By similarity).
CC -!- DISEASE: Defects in ED1 are the cause of the tabby phenotype in
CC mice (the equivalent of anhidrotic ectodermal dysplasia in
CC humans). The disease is characterized by sparse hair (atricichosis
CC or hypotrichosis), abnormal or missing teeth and the inability to
CC sweat due to the absence of sweat glands.
CC -!- SIMILARITY: Belongs to the tumor necrosis
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF016627; AAB95202.1; -
CC EMBL; AF016628; AAB95203.1; -
CC EMBL; AF016629; AAB95204.1; -
CC EMBL; AF016630; AAB95205.1; -
CC EMBL; AF016631; AAB95206.1; -
CC EMBL; AF004434; AAB88121.1; -
CC EMBL; AF004435; AAB88122.1; -
CC EMBL; Y13438; CAA73849.1; -
CC EMBL; AJ243657; CAB52696.1; -
CC EMBL; AJ243658; CAB52697.1; -
CC MGD; NGI1195272; Eda.
CC GO; GO:0045177; C:apical part of cell; IDA.
CC GO; GO:0005789; C:endoplasmic reticulum membrane; IDA.
CC GO; GO:0005887; C:integral to plasma membrane; IDA.
CC GO; GO:0007160; P:cell-matrix adhesion; IDA.
CC GO; GO:0042346; P:positive regulation of NF-kappaB protein-nu. . . ; IDA.
CC GO; GO:0007431; P:salivary gland development; IDA.
CC InterPro; IPR008160; Collagen.
CC InterPro; IPR006052; TNF family.
CC InterPro; IPR008983; TNF like.
CC Pfam; PF01391; Collagen; 1.
CC SMART; SM00207; TNF; 1.
CC PROSITE; PS00251; TNF_1; FALSE_NEG.
CC PROSITE; PS00049; TNF_2; 1.
KW Differentiation; Developmental protein; Collagen; Transmembrane;
KW Signal-anchor; Glycoprotein; Alternative splicing.
```





```
FT MOD_RES 665 665 METHYLATION (DI-).
FT MOD_RES 669 669 METHYLATION (DI-).
FT MOD_RES 673 673 METHYLATION (DI-).
FT MOD_RES 679 679 METHYLATION (DI-).
FT MOD_RES 681 681 METHYLATION (DI-).
FT MOD_RES 687 687 METHYLATION (DI-).
FT MOD_RES 691 691 METHYLATION (DI-).
FT MOD_RES 694 694 METHYLATION (DI-).
SQ SEQUENCE 713 AA; 76997 MW; 79DCDF724CED7DB4 CRC64;

Query Match 5.7%; Score 122.5; DB 1; Length 713;
Best Local Similarity 20.0%; Pred. No. 1.3;
Matches 53; Conservative 40; Mismatches 109; Indels 63; Gaps 8;

QY 98 KEYNALIDYPKKVGDLTDEEDDDGGGLDSTADDDDDSYSSVDDVGADYEDYTMLN 157
Db 133 KNGKNA-----KEDSDDEDDDDDDSDDEDEED-----EFEP----- 170
QY 158 KLNNAHTGTTPTSETTAEGETDSASSASNDNDVDFDFTSYNA-----HKKKQER-- 208
Db 171 PVVKGKQGVAAAPASEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 230
QY 209 ---KRSIADVNEEQNIQGNHTLOEKSNEATSK-----SPAPLHRRRMRHRRH 259
Db 231 PVKAKNVAEDDDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 290
QY 260 LLVRKARSEDSPAAHFLSLRRRHQSGMGVHGDMYICNDNERN-----YQCHQTR 312
Db 291 PEAKQKVEGESITPPFL-----FIGNLNPNKSVAEKLVASEPFAKN 334
QY 313 DGLVTVTNTGLYVYAQICYNNSHD 337
Db 335 DLAVDVVRTGTRKRGYVDFESAED 359

RESULT 12
TBP7_YEAST
ID TBP7_YEAST STANDARD; PRT; 1379 AA.
AC P40340;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE TAT-binding homolog 7.
GN YTA7 OR YGR270W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RA Agostoni Carbone M.L.; Lucchini G.; Melchiorretto P.; Nardese V.;
RA Vanoni M.; Panzeri L.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Feldmann H.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP PRELIMINARY SEQUENCE OF 1-1207 FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=95274317; PubMed=7754704;
RA Schnall R.; Mannhaupt G.; Stucka R.; Tauer R.; Ehnle S.;
RA Schwarzlose C.; Vetter I.; Feldmann H.;
RT "Identification of a set of yeast genes coding for a novel family of
RT putative ATPases with high similarity to constituents of the 26S
RT protease complex.";
RL Yeast 10:1141-1155(1994).
CC -!- SIMILARITY: Belongs to the AAA ATPase family.
CC -!- SIMILARITY: Contains 1 bromodomain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y07893; CAA69201.1; -.
DR EMBL; Z73055; CAA97300.1; -.
DR EMBL; X81072; CAA56963.1; -.
DR PIR; S64603; S64603.
DR GERMOnline; 141582; -.
DR SGD; S0003502; YTA7.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR003960; AAA_sub.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF00439; bromodomain; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00674; AAA; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
KW ATP-binding; Bromodomain.
FT DOMAIN 135 139 POLY-GLU.
FT DOMAIN 157 160 POLY-ARG.
FT DOMAIN 358 361 POLY-ASN.
FT DOMAIN 390 395 POLY-LYS.
FT NP_BIND 454 461 ATP (POTENTIAL).
FT DOMAIN 737 740 POLY-GLU.
FT DOMAIN 1044 1086 BROMODOMAIN (DIVERGENT).
FT CONFLICT 70 70 D -> E (IN REF. 2).
FT CONFLICT 241 241 S -> N (IN REF. 2).
FT CONFLICT 1016 1016 S -> N (IN REF. 2).
FT CONFLICT 1142 1142 S -> N (IN REF. 2).
FT CONFLICT 1153 1153 K -> E (IN REF. 2).
FT CONFLICT 1276 1276 I -> R (IN REF. 2).
FT CONFLICT 1283 1283 Q -> P (IN REF. 2).
SQ SEQUENCE 1379 AA; 157406 MW; 31D1F6F87E62E04F CRC64;

Query Match 5.7%; Score 122.5; DB 1; Length 1379;
Best Local Similarity 23.8%; Pred. No. 2.9;
Matches 53; Conservative 40; Mismatches 101; Indels 29; Gaps 9;

QY 87 INYLDEFDFEFOKEYENALIDYPK-KVQGLTDEEDDDGGGLDSTADDDDDSYSSVDDV 145
Db 44 INYAIKVFDFLEDDQVMDKDETPTVDTSDHEHNNKQK-----DDEDDVDLVSPHEN 98
QY 146 GADYEDYTDMLN-KLNNAHTGTTPTSETTAEGETDSASSASNDNDVDFDFTSYNAHK 204
Db 99 ARTNELTNERNLKRAKHA-----DPEEDDESFEEDVDDEDEDEDEFEDEYLEDSDKN 155
QY 205 KQERKRS-----IADVRNEEQNIQGNHTLOEKSNEATSKESAPLHRRMRHRRHL 260
Db 156 NRRRAADRAKRVFVDPDDDEYDEDD--EGDRIHSASSK-----RLKANSRR--- 203
QY 261 LVRKARSEDSPAAHFLSLRRRHQSGMGVHGDMYICNDNERN 303
Db 204 -TRSRHPETPPVRRALRSTRSRTSNEED--DENDNSRN 243

RESULT 13
OSTP_HUMAN
ID OSTP_HUMAN STANDARD; PRT; 314 AA.
AC P10451; Q15681; Q15682; Q15683;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Osteopontin precursor (Bone sialoprotein 1) (Urinary stone protein)
DE (Secreted phosphoprotein 1) (SPP-1) (Nephropontin) (Uropontin).
GN SPP1 OR OPN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89263749; PubMed=2726470;  
RA Kiefer M.C., Bauer D.M., Barr P.J.;  
RT "The cDNA and derived amino acid sequence for human osteopontin.";  
RL Nucleic Acids Res. 17:3306-3306(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90353945; PubMed=1974876;  
RA Young M.F., Kerr J.M., Termini J.D., Wever U.M., Wang M.G.,  
RA McBride O.W., Fisher L.W.;  
RT "cDNA cloning, mRNA distribution and heterogeneity, chromosomal  
RT location, and RFLP analysis of human osteopontin (OPN).";  
RL Genomics 7:491-502(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92108068; PubMed=1729712;  
RA Shiraga H., Min W., Vandusen W.J., Clayman M.D., Miner D.,  
RA Terrell C.H., Sherbotie J.R., Foreman J.W., Przysiecki C.,  
RA Neilson E.G., Hoyer J.R.;  
RT "Inhibition of calcium oxalate crystal growth in vitro by uropontin:  
RT another member of the aspartic acid-rich protein superfamily.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:426-430(1992).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Crosby A.H., Edwards S., Murray J.C., Dixon M.J.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=95031968; PubMed=7945249;  
RA Hijiya N., Setoguchi M., Matsuura K., Higuchi Y., Akizuki S.,  
RA Yamamoto S.;  
RT "Cloning and characterization of the human osteopontin gene and its  
RT promoter.";  
RL Biochem. J. 303:255-262(1994).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Yu W., Sarginson J., Gibbs R.A.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RX MEDLINE=95139605; PubMed=7637791;  
RA Saitoh Y., Kuratsu J., Takeshima H., Yamamoto S., Ushio Y.;  
RT "Expression of osteopontin in human glioma. Its correlation with the  
RT malignancy.";  
RL Lab. Invest. 72:55-63(1995).  
RN [8]  
RP SEQUENCE FROM N.A. (ISOFORM A).  
RC TISSUE=Kidney;  
RX MEDLINE=2238257; PubMed=12477932;  
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Altschuld S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max J., Wang J., Hsieh F.,  
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettenan M., Madan A.C., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[9]  
RP SEQUENCE OF 67-278 FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=92246977; PubMed=1575754;  
RA Kohri K., Suzuki Y., Yoshida K., Yamamoto K., Amaeaki N., Yamate T.,  
RA Umekawa T., Iguchi M., Sinohara H., Kurita T.;  
RT "Molecular cloning and sequencing of cDNA encoding urinary stone  
RT protein, which is identical to osteopontin.";  
RL Biochem. Biophys. Res. Commun. 184:859-864(1992).  
CC -!- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an  
CC integral part of the mineralized matrix. Probably important to  
CC cell-matrix interaction.  
CC -!- FUNCTION: Acts as a cytokine involved in enhancing production of  
CC interferon-gamma and interleukin-12 and reducing production of  
CC interleukin-10 and is essential in the pathway that leads to type  
CC I immunity (by similarity).  
CC -!- SUBUNIT: Ligand for integrin alpha-V/beta-3.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=A; Synonym=OPN-A, OP1B;  
CC IsoId=P10451-1; Sequence=Displayed;  
CC Name=B; Synonym=OPN-B, OP1A;  
CC IsoId=P10451-2; Sequence=VSP\_003778;  
CC Name=C; Synonym=OPN-C;  
CC IsoId=P10451-3; Sequence=VSP\_003777;  
CC -!- PTM: Extensively phosphorylated on serine residues.  
CC -!- PTM: N- and O-glycosylated.  
CC -!- DISEASE: This protein plays a principal role in urinary stone  
CC formation as the stone matrix.  
CC -!- SIMILARITY: Belongs to the osteopontin family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X13694; CAA31984.1; -  
CC EMBL; J04765; AAA59974.1; -  
CC EMBL; M83248; AAA17675.1; -  
CC EMBL; U20758; AAA68886.1; -  
CC EMBL; AF052124; AAC28619.1; -  
CC EMBL; D28759; BAA05949.1; -  
CC EMBL; D28760; BAA05950.1; -  
CC EMBL; D28761; BAA05951.1; -  
CC EMBL; D14813; BAA03554.1; -  
CC EMBL; BC017387; AAH17387.1; -  
CC PIR; S50028; S09575  
CC GlycoSuiteDB; P10451; -  
CC Genew; HGNC:11255; SPPI.  
CC MIM; 166490; -  
CC GO; GO:0005578; C:extracellular matrix; TAS.  
CC GO; GO:0008189; F:apoptosis inhibitor activity; ISS.  
CC GO; GO:0042056; F:chemotactant activity; TAS.  
CC GO; GO:0005225; F:cytokine activity; ISS.  
CC GO; GO:0003793; F:defense/immunity protein activity; TAS.  
CC GO; GO:0006953; F:growth factor activity; TAS.  
CC GO; GO:0005178; F:integrin binding; NAS.  
CC GO; GO:0006916; F:anti-apoptosis; ISS.  
CC GO; GO:0007267; P:cell-cell signaling; TAS.  
CC GO; GO:0007160; P:cell-matrix adhesion; NAS.  
CC GO; GO:0030595; P:immune cell chemotaxis; TAS.  
CC GO; GO:0006954; P:inflammatory response; TAS.  
CC GO; GO:0030502; P:negative regulation of bone mineralization; NAS.  
CC GO; GO:0042102; P:positive regulation of T-cell proliferation; TAS.  
CC GO; GO:0045637; P:regulation of myeloid blood cell differenti. .; TAS.  
CC GO; GO:0042088; P:T-helper 1 type immune response; TAS.  
CC InterPro; IPR002038; Osteopontin.









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:25:38 ; Search time 43 Seconds  
(without alignments)

3001.092 Million cell updates/sec

Title: US-09-813-329-6

Perfect score: 2162

Sequence: 1 MTATLKPFTTPTGANDGF.....DRNAVLREGNNRSYFGIFKV 409

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organalle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2162	100.0	409	5 Q8MY88	Q8MY88 drosophila
2	2157	99.8	409	5 Q8IGD3	Q8IGD3 drosophila
3	2149	99.4	415	5 Q8MUJ1	Q8MUJ1 drosophila
4	1656	76.6	325	5 Q9V5G2	Q9V5G2 drosophila
5	1359	62.9	261	5 Q8MRW2	Q8MRW2 drosophila
6	175.5	8.1	2226	5 Q97225	Q97225 plasmodium
7	172.5	8.0	2221	5 Q81259	Q81259 plasmodium
8	168	7.8	10061	5 Q81321	Q81321 plasmodium
9	165.5	7.7	3268	3 Q03280	Q03280 saccharomyc
10	164	7.6	2026	5 Q81565	Q81565 plasmodium
11	160.5	7.4	2309	5 Q81517	Q81517 plasmodium
12	159	7.4	491	5 Q81J07	Q81J07 plasmodium
13	154	7.1	2738	5 Q8IHU4	Q8IHU4 plasmodium
14	153.5	7.1	1371	5 Q8ICK4	Q8ICK4 plasmodium
15	153.5	7.1	3401	5 Q8IBQ3	Q8IBQ3 plasmodium
16	151	7.0	3452	5 Q8IEA3	Q8IEA3 plasmodium

#### ALIGNMENTS

RESULT 1

Q8MY88 ID Q8MY88 PRELIMINARY; PRT; 409 AA.  
AC Q8MY88;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE TNF superfamily ligand, Eiger (Tumor necrosis factor family member DTI).  
GN EIGER OR DT1 OR CG12919.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22060500; PubMed=12065414;  
RA Igaki T., Kanda H., Yamamoto-Goto Y., Kanuka H., Kuranaga E.,  
RA Aigaki T., Miura M.;  
RT "Eiger, a TNF superfamily ligand that triggers the Drosophila JNK pathway.";  
RT RT  
RL EMBO J. 21:3009-3018(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Inohara N., Nunez G.;  
RT "DTI, a Drosophila tumor necrosis factor family member.";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB073865; BAC00950.1; -  
DR EMBL; AF149799; AA015310.1; -  
DR Flybase; FBgn0033483; eiger.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.  
DR GO; GO:0006955; P:immune response; IEA.  
DR InterPro; IPR006052; TNF family.  
DR InterPro; IPR008983; TNF\_like.  
DR SMART; SM00207; TNF; 1.  
DR PROSITE; PS00251; TNF\_1; 1.  
DR PROSITE; PS50049; TNF\_2; 1.

Q8IM62 plasmodium  
Q8IE75 plasmodium  
Q9QZV4 mus musculus  
Q8I2R3 plasmodium  
Q8IID2 plasmodium  
Q9WVE4 mus musculus  
Q8IJQ2 plasmodium  
Q8I4Y8 plasmodium  
Q8I2500 saccharomyc  
Q8I298 plasmodium  
Q8IDY3 plasmodium  
Q8I1D5 plasmodium  
Q86EN6 schistosoma  
Q9U0L0 plasmodium  
Q8I350 plasmodium  
Q8ILB1 plasmodium  
Q8I1R1 plasmodium  
Q8MB1 dictyosteli  
Q8IM32 plasmodium  
Q8IKX1 plasmodium  
Q8I3R5 plasmodium  
Q86127 plasmodium  
Q8ILU2 plasmodium  
Q24376 drosophila  
Q61143 plasmodium  
Q8IE99 plasmodium  
Q8IKF6 plasmodium  
Q8I2K4 plasmodium

Query Match 99.8%; Score 2157; DB 5; Length 409;  
Best Local Similarity 99.8%; Pred. No. 3.1e-136;

```
Query Match      99.4%; Score 2149; DB 5; Length 415;
Best Local Similarity 98.6%; Pred. No. 1.1e-135;
Matches 409; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 MTAETLKPFITPTSANDGFFPAKATSTATAQRTTQLPLVLGFTGLGLVAILALTTWQ 60
Db 1 MTAETLKPFITPTSANDGFFPAKATSTATAQRTTQLPLVLGFTGLGLVAILALTTWQ 60

Qy 61 TTRVSHLDELKSLKRVVDNIQORIGINVLDFEFOKEYENALIDYPKKVDGLTDEDD 120
Db 61 TTRVSHLDELKSLKRVVDNIQORIGINVLDFEFOKEYENALIDYPKKVDGLTDEDD 120

Qy 121 DDGGLDSIADDEDDVSYSSVDDVGYEDYDMLNKLNNHAHTGTTTSETTAGEGET 180
Db 121 DDGGLDSIADDEDDVSYSSVDDVGYEDYDMLNKLNNHAHTGTTTSETTAGEGET 180

Qy 181 DSASASANDNDVDDFTSYNAHKKQKRSRSDVRNEEQNIQNHTELQKSSNEATS 240
Db 181 DSASASANDNDVDDFTSYNAHKKQKRSRSDVRNEEQNIQNHTELQKSSNEATS 240

Qy 241 KESAPLHRRMRHSHRHLVRK-----ARSEDSRPAAHFHLSSRRRHOGSMGYHGD 294
Db 241 KESAPLHRRMRHSHRHLVRK-----ARSEDSRPAAHFHLSSRRRHOGSMGYHGD 300

Qy 295 YIGNDNERNYSQGHFQTRDGLVLTNTGLYVYVYQIYNNNSHDQNGFVFGQDTPFLQCL 354
Db 301 YIGNDNERNYSQGHFQTRDGLVLTNTGLYVYVYQIYNNNSHDQNGFVFGQDTPFLQCL 360

Qy 355 NTVPTNMPHKVHTCHTSGLIHLERNERHLKDINHNRNAVLRGNNRSYFGIFKV 409
Db 361 NTVPTNMPHKVHTCHTSGLIHLERNERHLKDINHNRNAVLRGNNRSYFGIFKV 415

RESULT 4
QV5GZ PRELIMINARY; PRT; 325 AA.
AC QV5G2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE CG12919 protein.
GN EIGER OR CG12919.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=Berkeley.
EX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Epfnkoc C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
```

```
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreenek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY119233; RAMS1093.1; -.
DR FlyBase; FBgn0033483; elgcr.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR GO: GO:0006955; P:immune response; IEA.
DR InterPro: IPR006052; TNF family.
DR InterPro: IPR008983; TNF-like.
DR SMART; SM00207; TNF; 1.
DR PROSITE; PS00251; TNF 1; 1.
DR PROSITE; PS0049; TNF 2; 1.
DR PROSITE; PS0049; TNF 2; 1.
SQ SEQUENCE 261 AA; 29780 MW; 13B6D5A04EC9122C CRC64;

Query Match      62.9%; Score 1359; DB 5; Length 261;
Best Local Similarity 97.7%; Pred.No. 3.8e-83;
Matches 255; Conservative 0; Mismatches 0; Indels 6; Gaps 1,*
QY 155 MLNKLNNAAHTGTTPTSETTARCEGETDSASSASNDNDVDFDFTSYNAHKKKQERKRSIA 214
Db 1 MLNKLNNAAHTGTTPTSETTARCEGETDSASSASNDNDVDFDFTSYNAHKKKQERKRSIA 60
QY 215 DVNREEQNIQGNHTLQBSKSNKATSKESAPLHRRMRHSHRHLVLRK-----ARSE 268
Db 61 DVNREEQNIQGNHTLQBSKSNKATSKESAPLHRRMRHSHRHLVLRKGESLLSARSE 120
QY 269 DSRPAAPHLSRRRRHQSGMGVGHGDMYIGNDNERNYSQGHFQTRDGLVTWNTGLYVYVA 328
Db 121 DSRPAAPHLSRRRRHQSGMGVGHGDMYIGNDNERNYSQGHFQTRDGLVTWNTGLYVYVA 180
QY 329 QICYNNSHDQNGFIVFQGDTPFLQCLNTVPTNMPHKVHTCHTSLIHLERNERHLKDIH 388
Db 181 QICYNNSHDQNGFIVFQGDTPFLQCLNTVPTNMPHKVHTCHTSLIHLERNERHLKDIH 240
QY 389 NDRNAVLREGNRSYFGIFKV 409
Db 241 NDRNAVLREGNRSYFGIFKV 261

RESULT 6
O97225 PRELIMINARY; PRT; 2226 AA.
AC O97225;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein, conserved.
GN MAL3P2.2.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99376085; PubMed=10448855;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
RA Gentles S., Gilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moule S.,
RA Mungall K., Murphy J., Oliver K., Quail M.A., Rajandream M.-A.,
RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
RT falciparum .";
RL Nature 400:532-538 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
```

```
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531 (2002).
DR EMBL: AL034558; CAB38989.3; -.
DR InterPro: IPR007259; SPC97_Spc98.
DR Pfam; PF04130; SPC97_Spc98; 1.
KW Hypothetical protein.
SQ SEQUENCE 2226 AA; 267974 MW; 8690501ED4994768 CRC64;

Query Match      8.1%; Score 175.5; DB 5; Length 2226;
Best Local Similarity 21.9%; Pred.No. 0.0062;
Matches 82; Conservative 57; Mismatches 150; Indels 85; Gaps 16;
QY 69 KELKSLKRVVDNLQORLGINYLDEFDEF-----QKEYENALIDYPKKVDGLTDEEDDDG 123
Db 155 KETKRRKRTLHNKN-----DTFGSVTSMNDEKDYMKYNQDYHDKDDDDYDEDDDED 207
QY 124 DGLDSTADDEDDVSYSVDVGDYEDYDMLKLNNAHTGTTPTSETTARCEGETDSA 183
Db 208 D--DDYDDEDDDD--YEDDDDD--DDYDDEDDNNNSYDQNTKTKKKINPM 258
QY 184 SSASNDNDNVDFDFTSYNAHKKKQERKRSIADVRN-----EEQNIQGNHTLQBSK 232
Db 259 YNSTFETSNMFCVKDEKKEKQKNSI--DKRNIIYDSDSDNNYEHIFTHNSDLFF 317
QY 233 -KSSNEATSKESAPLHRRMRHSHRHLVLRKARSEDSRPAAPHLSRRRHQSGMGVH 291
Db 318 LSIISNDHIEKEN--NLYIQNEQFINYDVFIRKPK-----FH 352
QY 292 GDMYIGNDNER-----NSYOGHFQTRDGLVTWNTGLYVY--YAQI--CYNNS 335
Db 353 KNLVFNMDSCNSWIVKKSILINEITRSQKND--LVINDVLYDVNVNHYRINYCYTRL 410
QY 336 HDQNGFIVFQGDTPFL-----QCLNTVPTNMPHKVHTCHTSLIHLERNERHLKDIH 388
Db 411 YINDDFLLNNKTKELGVQNVASHINNTNNNNNNIDNIHNVNNIN--NYNWREHLKRIE 469
QY 389 NDRNAVLREGNRS 402
Db 470 KKKKNNNNNNNN 483

RESULT 7
Q81259 PRELIMINARY; PRT; 2221 AA.
ID Q81259;
AC Q81259;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PFA0410W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
```

RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,  
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,  
RA Cronin A., Davies R., Davis P., Dearden F., Doggett J.,  
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,  
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,  
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,  
RA Knights A., Konfortov B., Kyes S., Larkie N., Lawson D., Lennard N.,  
RA Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,  
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,  
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,  
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,  
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,  
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;  
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.:"  
RL Nature 419:527-531(2002).  
DR EMBL; AL031745; CAD49042.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 221 AA; 255042 MW; 7311E1373C1B6844 CRC64;  
  
Query Match 8.0%; Score 172.5; DB 5; Length 2221;  
Best Local Similarity 21.2%; Pred. No. 0.0099;  
Matches 83; Conservative 61; Mismatches 153; Indels 95; Gaps 15;  
  
QY 67 LDKEKLSI-----KRVVDNL-----QQRIG--INYLDFEFQKEYENALIDYPKKVDGLTD 116  
Db 261 LDKNEKSNFNDKNADENMIKLDSSKLGKNGKFKNENDEFEEBEIDVD---DVDDVDVD 317  
QY 117 EDDDDGGLGSLADDEDDDDSYSSVDVGVADYEDYDTMLKLNNAHTGTTPTSETTAE 176  
Db 318 VDDVDVDDVDVDDDE 370  
QY 177 EGETDSASSANDNVF-----DFTSYNAHKKQ-----ERKRSIADVRNEEQINQNH 227  
Db 371 EDDDDDEYDDDDDEYDE 430  
QY 228 TELQKESNEATSKESPAPLHRRHMRHSHRHLVRKARSDSRPAAPHLSRRRRHOGS 287  
Db 431 EMVVENKKNINNEDEKNIIDKKSMNKKR-----KKKKKRV 468  
QY 288 MGYHGDVIGNDNERNSSYGHFQ--TRDGLVTVNTGLVYVYQIYCNNSHDQNGFI--- 342  
Db 469 NNN 522  
QY 343 ---VFQGTPTFLQCLNTVPTNMP-HKVHTCHT-----SGLIHLERNE 380  
Db 523 ASDMYNKDTN-MSSLINTSENTLPVLIILDCSAVLDMKELWKDKSLTLPSPFGLLYLNK 581  
QY 381 -----RIHLKDIHNDRAVLREGNRS 402  
Db 582 LLKANTIGNNNIHMNN 613  
  
RESULT 8  
Q81321 PRELIMINARY; PRT; 10061 AA.  
AC Q81321;  
ID Q81321;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN PF05070W.  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,  
RA Hall N., Bowman S., Churcher C., Quail M., Barrell B.;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22255708; PubMed=12368867;  
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,

RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,  
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,  
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,  
RA Cronin A., Davies R., Davis P., Dearden F., Doggett J.,  
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,  
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,  
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,  
RA Knights A., Konfortov B., Kyes S., Larkie N., Lawson D., Lennard N.,  
RA Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,  
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,  
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,  
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,  
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,  
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;  
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.:"  
RL Nature 419:527-531(2002).  
DR EMBL; AL929352; CAD51479.1; -.  
DR GO; GO:000509; F:calcium ion binding; IEA.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0008318; F:protein prenyltransferase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0018346; P:protein amino acid prenylation; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR002048; EF-hand.  
DR InterPro; IPR002083; MATH.  
DR InterPro; IPR002088; PPTA.  
DR InterPro; IPR006145; PseudoU\_synth.  
DR Pfam; PF00917; MATH; 1.  
DR Pfam; PF00849; PseudoU\_synth\_2; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00018; EF\_HAND; 1.  
DR PROSITE; PS00144; MATH; 1.  
DR PROSITE; PS00904; PPTA; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 10061 AA; 1187600 MW; F13CE7E4FE86F3DD CRC64;  
  
Query Match 7.8%; Score 168; DB 5; Length 10061;  
Best Local Similarity 20.8%; Pred. No. 0.12;  
Matches 86; Conservative 80; Mismatches 127; Indels 120; Gaps 22;  
  
QY 66 HLDKELSLKRVVDNLQORLGINYLDFEFQKEYENALIDYPKKVDGLTDEEDDDG 125  
Db 8288 YLDEEGKVSXK--KMKKKLLN--DK--EHEKDNEDNEDNEDDEDEDEDEDEDEDE 8340  
QY 126 LDSADDEDDDDSYSSVDVGVADY-EDYDTML-----NKLNAHTGTTPTSETTAE 179  
Db 8341 DDDDDDDDDDD-----DDYDEYDEYDEYDEYDEYDEYDEYDEYDEYDEYDEYDE 8393  
QY 180 TDSASSASNDNVDFDFTSYNAHKKQER-----KRSIADV-----RNEE--- 220  
Db 8394 TNNRSSSNTIDMKKTIYEKNCMKKEYRTHDLITNESIMYKNEIYDKEKYRQENIN 8453  
QY 221 -----QNIQGNH--TELOEK-----SSNEATSKES-----PAPLHH 249  
Db 8454 HMSNMYMDKNINHHEHYDNNKKKNYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 8513  
QY 250 RRMHSRHRLLVRKARSDSRPAAPHLSRRRRHOGSMGYHGDVIGNDNERNSS 304  
Db 8514 NVKE-----TEKNVQQNL-----PHISKNNNN-----NNDSNTNNSYDY 8550  
QY 305 --YQGHFQTRDGLTVTNTGLY-----VYAQICYNNSHDQNGFIVFOGDTPLQLCLNT 356  
Db 8551 NMSNSNYGSKMGSGNNINDSTNYNNMREKNIYNIILCNN--DSNNYVLFNSNEKY----- 8601  
QY 357 VPTNMPHKVHTCHTSGLI-HLERNRIHL-KDINHNDRAVLREGNRSYGF 407  
Db 8602 ---SMNKNISNLSNMMKNQDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 8651  
  
RESULT 9  
Q03280  
ID Q03280  
AC Q03280;  
PRELIMINARY; PRT; 3268 AA.



RESULT 11	Q81517	PRELIMINARY;	PRT; 2309 AA.
ID	Q81517		
AC	Q81517		
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)	
DE	Hypothetical protein.		
GN	PFL1980C.		
OS	Plasmodium falciparum (isolate 3D7).		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=36329;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=3D7;		
RX	MEDLINE=22255705; PubMed=12368864;		
RA	Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,		
RA	Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,		
RA	Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,		
RA	Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,		
RA	Pertea M., Allen J., Salengut J., Haft D., Mather M.W., Vaidya A.B.,		
RA	Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,		
RA	McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,		
RA	Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,		
RA	Fraser C.M., Barrell B.;		
RT	"Genome sequence of the human malaria parasite Plasmodium		
RT	falciparum.";		
RL	Nature 419:498-511(2002).		
DR	EMBL; AE014850; AAN36480.1; --		
DR	Hypothetical protein.		
SK	SEQUENCE 2309 AA; 276528 MW; 8AGF8DBB972C4922 CRC64;		
Query Match	7.4%; Score 160.5; DB 5; Length 2309;		
Best Local Similarity	19.6%; Pred. No. 0.066;		
Matches	70; Conservative 77; Mismatches 118; Indels 93; Gaps 15;		
Qy	69	KELKSLKRVVDNLQORLGINYLDFD--EFOKEYENALIDYPKVDGLTDEDDDDGDL	126
Db	701	KYLMKYKNVSNLQ-----DFNKDYSESEGEVNDPVKL-----DQSNINRGS	747
Qy	127	DSIADDEDDVYSVSDVGA-DYEDYD-----MLKLNNAHTGTPPTSETAEGE	179
Db	748	YTSHYNKGVDVNTSSYDKNSYSDDDYDTSKFQENRYKHSYK-NVYKNKSSDDYEND	806
Qy	180	TDSASSASNDNVF--DFTSYNAHKQKQERKSIADVRNEEQINQNHTELEKS---	234
Db	807	SNSSNNNNNTYSDDDFNSYNNRRNTQKSKQMKNERQKKNKKNKIDQSDSTMSD	866
Qy	235	-----SNEATSKESPAPLHRRHRMHSRHLVRKARSEDSPAAPHFLSSRRHQ	285
Db	867	EYNNMNDLYSNEYTNRSIFNDILHAKKRFYDLISINRMKNT-----	911
Qy	286	GSMGYHGDYMGINDNERNYSOGHFQTRDGVLTVTNTGLYVYVAOICYNNSHDQGFIVQ	345
Db	912	-----FNVNNINKNTID--ENYIENIDTIS---YFT---SIYNSEH-----IFQ	948
Qy	346	GDPFLOCLNTVPTNMPKHVHTCHTSGLIHLERNERHLKDIHNDRAVLREGNRSY	403
Db	949	FDYFLFAIEICTYNNYI-----KSYLINNNNN---NNNNLKY	988
RESULT 12			
Q81JQ7			
ID	Q81JQ7	PRELIMINARY;	PRT; 491 AA.
AC	Q81JQ7		
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)	
DE	Hypothetical protein.		
GN	Pf10_0135.		
OS	Plasmodium falciparum (isolate 3D7).		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 02:20:57 ; Search time 6712 Seconds  
(without alignments)  
2641.135 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTAEFLKPFITPTSANDGF.....DRNAVLREGNNRSYGFIFKV 409

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09813329/runat\_13092004\_102753\_21667/app\_query.fasta\_1.583  
-DB=genEmbl -QWMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09813329 -CGN 1\_1 7406 @runat\_13092004\_102753\_21667 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.hcg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.ats.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.ats.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2162	100.0	1656	3	AB073865	AB073865 Drosophil
2	2162	100.0	2101	3	AF149799	AF149799 Drosophil
3	2157	99.8	2165	3	BT001838	BT001838 Drosophil
4	2149	99.4	1248	3	AY115551	AY115551 Drosophil
5	2149	99.4	2159	3	AF521176	AF521176 Drosophil
6	1626	75.2	1221	3	AY119233	AY119233 Drosophil
7	1508.5	69.8	55359	2	AC012854	AC012854 Drosophil
c 8	1508.5	69.8	158983	3	AC005974	AC005974 Drosophil
c 9	1508.5	69.8	172904	3	AC007414	AC007414 Drosophil
c 10	1508.5	69.8	189620	3	AC099029	AC099029 Drosophil
c 11	1508.5	69.8	275390	3	AE003831	AE003831 Drosophil
12	175.5	8.1	158548	3	PFMAL3P2	AL034558 Plasmodiu
13	172.5	8.0	110000	3	PFMAL1P2_2	Continuation (3 of
14	168	7.8	313050	3	PFA329352	AL929352 Plasmodiu
15	165.5	7.7	69009	8	SCD8035	U33050 Saccharomyc
16	164	7.6	196490	2	AC005507	AC005507 Plasmodiu
c 17	164	7.6	250421	3	AE014849	AE014849 Plasmodiu
c 18	160.5	7.4	250713	3	AE014850	AE014850 Plasmodiu
c 19	160.5	7.4	256172	2	AC005139	AC005139 Plasmodiu
c 20	160.5	7.4	310779	2	AC005140	AC005140 Plasmodiu
c 21	154	7.1	110000	2	PFMAL13_08	Continuation (9 of
c 22	154	7.1	253151	3	AE014842	AE014842 Plasmodiu
c 23	154	7.1	318221	2	PFMAL13P3	AL049184 Plasmodiu
24	153.5	7.1	2810	4	RABHRCB	JO5080 Rabbit hist
25	153.5	7.1	110000	2	PFMAL6P1_07	Continuation (8 of
26	153.5	7.1	110000	2	PFMAL7P1_07	Continuation (8 of
c 27	152	7.0	267206	2	AC131871	AC131871 Rattus no
c 28	152	7.0	269285	2	AC120096	AC120096 Rattus no
c 29	151.5	7.0	2270	10	AF158597	AF158597 Mus muscu
c 30	151.5	7.0	133661	14	U93872	U93872 Kaposi's sa
c 31	151.5	7.0	208632	2	AC073711	AC073711 Mus muscu
c 32	151	7.0	565	6	BD231127	BD231127 Shuffling
33	151	7.0	2616	6	AX642183	AX642183 Sequence
34	150.5	7.0	250029	3	AE014816	AE014816 Plasmodiu
35	149	6.9	5301	9	AF061189	AF061189 Homo eapi
c 36	148.5	6.9	110000	2	PFMAL13_09	Continuation (10 o
c 37	148.5	6.9	110000	2	PFMAL13_10	Continuation (11 o
c 38	148	6.8	250029	3	AE014839	AE014839 Plasmodiu
39	148	6.8	341050	3	PFA929357	AL929357 Plasmodiu
c 40	147.5	6.8	251762	3	AE014851	AE014851 Plasmodiu
c 41	147	6.8	5057	8	SCYLR114C	273286 S.cerevisia
c 42	147	6.8	15693	8	YSCY354	U53878 Saccharomyc
c 43	147	6.8	37639	8	SCCHX11LA	X89514 S.cerevisia
c 44	146	6.8	2407	10	AF132218	AF132218 Mus muscu
45	146	6.8	250195	3	AE014831	AE014831 Plasmodiu

ALIGNMENTS

RESULT 1

AB073865	LOCUS	AB073865	1656 bp	mRNA	linear	INV 27-JUN-2002
DEFINITION	Drosophila melanogaster mRNA for TNF superfamily ligand, Eiger, complete cds.					
ACCESSION	AB073865					
VERSION	AB073865.1	GI:21623741				
KEYWORDS						
SOURCE	Drosophila melanogaster (fruit fly)					
ORGANISM	Drosophila melanogaster					
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Epiphytoidea; Drosophilidae; Drosophila.					
REFERENCE	1	Igaki,T., Kanda,H., Yamamoto-Goto,Y., Kanuka,H., Kuranaga,E., Aigaki,T. and Miura,M. Eiger, a TNF superfamily ligand that triggers the Drosophila JNK pathway EMBO J. 21 (12), 3009-3018 (2002)				
AUTHORS						
TITLE						
JOURNAL						
PUBMED	12065414					
REFERENCE	2	(bases 1 to 1656)				
AUTHORS	Igaki,T.					
TITLE	Direct Submission					
JOURNAL	Submitted (02-NOV-2001) Tatsushi Igaki, Brain Science Institute, RIKEN, Laboratory for Cell Recovery Mechanisms; 2-1 Hirosawa, Wako, Saitama 351-0198, Japan (E-mail: igaki@brain.riken.go.jp, Tel:81-48-467-6945, Fax:81-48-467-6946)					
FEATURES	Location/Qualifiers					
source	1..1656	/organism="Drosophila melanogaster"				
		/mol_type="mRNA"				
		/db_xref="taxon:7227"				
gene	1..1656					
		/gene="eiger"				
CDS	427..1656					
		/gene="eiger"				
		/codon_start=1				
		/evidence=experimental				
		/product="TNF superfamily ligand, Eiger"				
		/protein_id="BAC00950.1"				
		/db_xref="GI:21623742"				
		/translation="MTAETLKPIPTTSANDDGFPKAKATSTATAQRTRQLPLVLGFGIGLVVALIALIIVQTTVRSHLDKLKSIRVDNLQQRLGINLIDEEDEFOKEYENALDIYPRKVGDLIDEEDDDGGDLSADDEDVVSDVDDGVADIEDYTDMLNKINNAHTQTPTSETTAAGEGEATSASSASNDNVEDFTSYNAHKKKQEKRSRGIADVNEEQNIQNHTTLEAGSENEATSKESPAPLHRRHRMHSRHLLIVRKARSDSRPAAHFLHSRRHQSGMGVGHGYTGNDNRNSYQGHPFRDGLVTVTNTGLYVYVAQICVNNSPHQNGPFI VFGDTPTPLQCLNTVPTNMHPKVHTCHTSLGILHLERNERIHLKOLHNDRN AVLREGNRSRYFGIPKV"				
ORIGIN						
Alignment Scores:						
Pred. No.:	2, 02e-142	Length:	1656			
Score:	2162.00	Matches:	409			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	3	Gaps:	0			
US-09-813-329-6	(1-409) x AB073865 (1-1656)					
QY	1	MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe	20			
Db	427	ATGACTGCCGAGACCTTCAGCGCTTTATTACCACGAGTGCCACCATGATGTTT	486			
QY	21	ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrGlnLeulleProLeu	40			
Db	487	CCGCGCCAAGCGCACCAGCACGGCGACCGCCAGCGCACCGCATGCCCTTG	546			
QY	41	ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleAlaIleuThrIleTrpGln	60			
Db	547	GTITTCGGGTTCATCGGTCTGGGGCTCGTGCCTTCCTCGCATCTCGCATCAACGATCGGCAG	606			
QY	61	ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn	80			

**LOCUS** AF149799 2101 bp mRNA linear INV 02-JAN-2003  
**DEFINITION** Drosophila melanogaster tumor necrosis factor family member DT1  
**ACCESSION** AF149799  
**VERSION** AF149799.1 GI:27462085  
**KEYWORDS**  
**SOURCE** Drosophila melanogaster (fruit fly)  
**ORGANISM** Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 2101)  
**REFERENCE**  
**AUTHORS** Inohara, N. and Nunez, G.  
**TITLE** DT1, a Drosophila tumor necrosis factor family member  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 2101)  
**AUTHORS** Inohara, N. and Nunez, G.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (10-MAY-1999) Department of Pathology, Comprehensive  
Cancer Center, 1500 E. Medical Center Dr., Ann Arbor, MI 48109, USA  
**FEATURES**  
**source**  
1..2101  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
1..2101  
/gene="dt1"  
625..1854  
/gene="dt1"  
/codon\_start=1  
/product="tumor necrosis factor family member DT1"  
/protein\_id="AAO15310.1"  
/db\_xref="GI:27462086"  
/translation="MTAEELKPFITPTSDNDGPPAKATSPATAQRETROLIPVLGF  
IGLGIVAILALTWQTRVSHLDKELSKLVVDNLQORLGINLYDEFDFQKEYEN  
ALIDYPKKVDGTEDDDDGLDSDADDDDDVSYSSVDVGDVYEDYDMLKLN  
NAHTGTTPTSETTAEGGETDSASASNDNDVDFDFTSYNAHKKKQKRSRSTADVRN  
EEQNIQGNHTLQKSNTEATSKESPAPLHRRMRHSHRHLVLRKARSDSRPAAHF  
HLSRRHQGSMGVHDMYIGNDNERNYSQGHFOTRDGVLTVNTGLVYVVAQICYN  
SHDQNGFIVFGDTPFLQCLNTVPTNNPKHVHTCHTSLIHLERNERHLKDIHNDRN  
AVLREGNRSYFGIEKV"  
2080..2085  
/gene="dt1"  
2101  
/gene="dt1"  
**ORIGIN**  
Alignment Scores:  
Pred. No.: 2,65e-142 Length: 2101  
Score: 2162.00 Matches: 409  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0  
US-09-813-329-6 (1-409) x AF149799 (1-2101)  
Qy 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe 20  
Db 625 ATGACTGCCGAGACCCCTCAACCGCTTTATACGCCCAACGAGTGCACGATGATGGTTT 684  
Qy 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40  
Db 685 CCGGCCAAGGACGACGACGCGACCCGACGCGACGCGACCCGCGCTGATCCCCCTG 744  
Qy 41 ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTrpGln 60  
Db 745 GTTTTGGGGTTCATCGGCTCGGGCTGGTTCGTCATTCTCGCACTAACCATCTGGCAG 804  
Qy 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn 80  
Db 805 ACAACGCGGTGATCGCATCTCGCAAGGAGCTGAAGAGCCTGAAGCGAGTCTGCATAAT 864

Qy 81 LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100  
Db 865 CTCAGCAGCGTTTGGGCATAAACTATCTGACGAGTTCGACGAGTTCCTCAAAAGAGTAC 924  
Qy 101 GluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluGluAspAsp 120  
Db 925 GAGAATGCCCTCATCGACTATCCAAAAAAGGTGGATGGCTCTACGATGAGGAGGACGAC 984  
Qy 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer 140  
Db 985 GACATGGCATGCTCTGGATTCCATTGCGACGACGAGGACGACGAGCTTAGCTATAGC 1044  
Qy 141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn 160  
Db 1045 TCTGTGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAATAAATACTCAAC 1104  
Qy 161 AsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180  
Db 1105 AATGCACATACCGGCACCCACCATCTGAGACCACTCTGAGGCGGAGGCGAGACG 1164  
Qy 181 AspSerAlaSerSerAlaSerAsnAspAspValPheAspAspPheThrSerTyrAsn 200  
Db 1165 GACATGCATCTCTCAGCTCAATATGATGACATGTGTTCGATGACTTTTACCGACTACAT 1224  
Qy 201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu 220  
Db 1225 GCCCACAAAAAAGCAGGAGGAGAGAAAATCTCGCTCGATTGCCGATGTACGCAATGAGGAG 1284  
Qy 221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlySerSerAsnGluAlaThrSer 240  
Db 1285 CAGATAATTCAGGAAATCACACAGAGCTTCAGGAAAGATCATCCAATGAGGCAACTTCC 1344  
Qy 241 LysGluSerProAlaProLeuHisLeuArgArgArgMetHisSerArgHisLeuGluLeu 260  
Db 1345 AAGAGAGACCCCTGCACCACTTACCACCGTCGAGATGATTTCCGCCCATCGCACCTC 1404  
Qy 261 LeuValArgLysAlaArgSerGluAspSerArgProAlaHisPheHisLeuSerSer 280  
Db 1405 CTAGTCGCAAAAGCCAGATCCGAGGACTCGAGGCGCAGCAGCCCACTTTCCACTTCGAGCAGC 1464  
Qy 281 ArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn 300  
Db 1465 AGCGCGGTCAACCAAGGAGTATGGGCTACCATGGAGATGTACATAGGAAATGATAAC 1524  
Qy 301 GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn 320  
Db 1525 GAGAGAAACTCTTATCAGGACACTTTCAAACCGCGCATGGCTCTTGACGGTGACCAAT 1584  
Qy 321 ThrGlyLeuTyrTyrValTyrAlaGlnIleCysThrAsnAsnSerHisAspGlnAsnGly 340  
Db 1585 ACAGGCCTATATTACGTATACGCCCATGATATGCTACAACTCGCAGCAGCAGACGGA 1644  
Qy 341 PheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsn 360  
Db 1645 TTTATCGTCTTCAAAGGAGACACTCCATTCTCGAGTGTGTAACACGCGTCCCAAC 1704  
Qy 361 MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGlu 380  
Db 1705 ATGCCAATAAGGTGCACACTGCCACACGAGTGGTCTGATCCACTTGGAAACGAAACGAG 1764  
Qy 381 ArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsn 400  
Db 1765 AGGATCCATCTGAAGGACATTCACACGATCGCAATGCACTTCTTCGCGGGAGGAAACAAC 1824  
Qy 401 ArgSerTyrPheGlyIlePheLysVal 409  
Db 1825 CGAAGCTACTTTGGCATCTTCAAGGTG 1851  
**RESULT 3**  
BT001838  
LOCUS BT001838 2165 bp mRNA linear INV 15-NOV-2002  
DEFINITION Drosophila melanogaster RH51659 full insert cDNA.  
ACCESSION BT001838



QY 381 ArgileHisLeuLysAspIleHisAenAspArgAenAlaValLeuArgLysGluGlyAenAen 400  
 |||||  
 Db 1788 AGGATCCATCTGAAGGACATTACAAACGATCGCATGCGAGTCTCTGCGGGAGGAAACAC 1847  
 |||||  
 QY 401 ArgSerTyrPheGlyIlePheLysVal 409  
 |||||  
 Db 1848 CGAAGCTACTTTGGCATCTTCAAGGTG 1874  
 |||||

RESULT 4  
 AY115551

LOCUS AY115551 1248 bp mRNA linear INV 01-AUG-2003  
 DEFINITION Drosophila melanogaster DARTH (darth) mRNA, complete cds.  
 ACCESSION AY115551  
 VERSION AY115551.1 GI:31321973

KEYWORDS  
 SOURCE  
 ORGANISM  
 Drosophila melanogaster (fruit fly)  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 1248)  
 Kauppila,S., Maaty,W.S., Chen,P., Tomar,R.S., Eby,M.T., Chapo,J.,  
 Chew,S., Rathore,N., Zachariah,S., Sinha,S.K., Abrams,J.M. and  
 Chaudhary,P.M.  
 Eiger and its receptor, Wengen, comprise a TNF-like system in  
 Drosophila

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 2275938  
 12894227

REFERENCE 2 (bases 1 to 1248)  
 Chaudhary,P.M.  
 Direct Submission  
 TITLE  
 JOURNAL  
 Submitted (28-MAY-2002) Internal Medicine, UT Southwestern Medical  
 Center, 5323 Harry Hines Blvd., Dallas, TX 75390, USA

FEATURES  
 source  
 1..1248  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"  
 /chromosome="2"  
 /map="2R"  
 1..1248  
 /gene="darth"  
 1..1248  
 /gene="darth"  
 /note="Drosophila aspartate rich TNF homolog; tumor  
 necrosis factor ligand family member"  
 /codon\_start=1  
 /product="DARTH"  
 /protein\_id="AAM66763.1"  
 /db\_xref="GI:31321974"  
 /translation="MTAETLKPFTTPTSANDDGPAPKATSTATAQRTRQLIPLVLGF  
 IGLGLVALIALTITQTRVSHLDKELSKRVVDNIQQRLGINYLDFEFQKEYEN  
 ALIDYPRKVDGLTDEEDDGDGLDSIADDEDDVSYSVDDVGADYEDYDMLKLN  
 NAHTPTPTSETTAEGGETDSASANDNVDFDFTSYNAHKKKQKRSIAADVRN  
 EEQNIQGNHTLEQKSNEATSKESPAPLHRRMRHSRHRLLVVKGESLLSARSEDS  
 RPAAHFLHSRSPRQSGNGHYGDMYIGNDNERNYSYQGHFQTRDGLVLTWTGLYYVA  
 QICYNNSHDQNGFVFGDTPFLQCLNTVPTNMPKHVHTCHTSLGHLHNERIHLKD  
 IHNDRNAVLRGNRSYFGIPKV"  
 1..108  
 /gene="darth"  
 /note="Region: intracellular domain"  
 109..177  
 /gene="darth"  
 /note="Region: transmembrane domain"  
 178..1245  
 /gene="darth"  
 /note="Region: extracellular domain"  
 433..438  
 /gene="darth"  
 /note="cleavage site"

misc\_feature  
 1..108  
 /gene="darth"  
 /note="Region: intracellular domain"  
 109..177  
 /gene="darth"  
 /note="Region: transmembrane domain"  
 178..1245  
 /gene="darth"  
 /note="Region: extracellular domain"  
 433..438  
 /gene="darth"  
 /note="cleavage site"

ORIGIN

Alignment Scores:  
 Pred. No.: 1.19e-141 Length: 1248  
 Score: 2149.00 Matches: 409  
 Percent Similarity: 98.55% Conservatives: 0  
 Best Local Similarity: 98.55% Mismatches: 0  
 Query Match: 99.40% Indels: 6  
 DB: 3 Gaps: 1

US-09-813-329-6 (1-409) x AY115551 (1-1248)

QY 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAenAspAspGlyPhe 20  
 |||||  
 Db 1 ATGACTCCGAGACCCCTCAAGCCGCTTTATAACGCCAACGAGTGCACACGATGATGGTGT 60  
 |||||  
 QY 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuLysProLeu 40  
 |||||  
 Db 61 CCGGCCAAAGCGACCGACGCGCGCCGACGCGACGCCGCCAGCTGATCCCCCTG 120  
 |||||  
 QY 41 ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTyrGln 60  
 |||||  
 Db 121 GTTTTGGGGTTCATCGGTCTGGGGCTGGTTCGTTGCCATTCTCGCACTAACGATCTG 180  
 |||||  
 QY 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAen 80  
 |||||  
 Db 181 ACAACGGGTGTATCGCATCTGGACAAGGAGCTGAAGAGCCCTGAAGCGAGTCTCGATA 240  
 |||||  
 QY 81 LeuGlnGlnArgLeuGlyIleAenTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100  
 |||||  
 Db 241 CTCGACGAGCGTTTGGGCATAAACTATCTGGACGAGTTCGACGAGTTCCTCAAAAGGAG 300  
 |||||  
 QY 101 GluAenAlaLeuLysAspTyrProLysLysValAspGlyLeuThrAspGluGluAspA 120  
 |||||  
 Db 301 GAGATCGCTCATCGCTATCCAAAAGAGTGGATGGCTCAGCGATGAGGAGACGAC 360  
 |||||  
 QY 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer 140  
 |||||  
 Db 361 GACGATCGCATGCTCGGATTCATTTCGCGACGACGAGGACGACGAGTTCGATATAGC 420  
 |||||  
 QY 141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAenLysLeuA 160  
 |||||  
 Db 421 TCTGTGGATGATGTGGCGCAGACTACGAGGACTACACCGATGTGTTAAATAAACA 480  
 |||||  
 QY 161 AenAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyThr 180  
 |||||  
 Db 481 AATGCATATCCGGCACACGCCACATCTTGAGACCCTCTGAGGGGGGAGACG 540  
 |||||  
 QY 181 AspSerAlaSerSerAlaSerAenAspAspValPheAspAspPheThrSerTyrAen 200  
 |||||  
 Db 541 GACAGTGCATCCTCAGCCTCAAATGATGACAATGTGTTCGATGACTTTACCGACTACA 600  
 |||||  
 QY 201 AlaHisLysLysLysGlnLysSerArgSerIleAlaAspValArgAenGluGlu 220  
 |||||  
 Db 601 GCCCACAAAAGAAGCAGGAGAGAAAATCTCGTTCGATGCCGATGATGACCAATGAG 660  
 |||||  
 QY 221 GlnAenIleGlnLysAenHisThrGluLeuGlnLysSerSerAenGluAlaThrSer 240  
 |||||  
 Db 661 CAGATATTTCAGGAATATCACAGAGCTTCAGAAAGATCTCCATTCATGAGGCACTTC 720  
 |||||  
 QY 241 LysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHisLeu 260  
 |||||  
 Db 721 AAAGAGAGCCCTGCACCTTCCACCCTCGCAGATGATGATTCCTCCGCCATCGCCACTC 780  
 |||||  
 QY 261 LeuValArgLys-----AlaArgSerGluAspSerArgProAlaAla 274  
 |||||  
 Db 781 CTAGTCCGCAAGGTGAATCTCTTCTTCAGCCAGATCCCGAGGACTCGAGGCCAGCACC 840  
 |||||  
 QY 275 HisPheHisLeuSerSerArgArgHisGlnGlySerMetGlyTyrHisArgLysMet 294  
 |||||  
 Db 841 CATTTCCATTGAGCAGCAGCGCGCTCACCAAGGAAGTATGGGCTACCATCGAGATATG 900  
 |||||  
 QY 295 TyrIleGlyAenAspAenGluArgAenSerTyrGlnGlyHisPheGlnThrArgAspGly 314  
 |||||



Db	901	TACATAGGAATGATAACAGAGAACTCTTATCAGGACACTTTCAAACGGCGATGCC	960
QY	315	ValLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsn	334
Db	961	GTCTTGACGGTACCAATACAGCCCTATATTACGTATACGCCAGATGCTTACAACAAC	1020
QY	335	SerHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeu	354
Db	1021	TGCGACGACAGAGCGATTTATCGTCTTTCAAGGAGACACTCCCAATTCCTGCAGTGTG	1080
QY	355	AsnThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIle	374
Db	1081	AACACGGTGCACCAACATGCCATAGGTGCACACTGCCACACAGTGTCTGATC	1140
QY	375	HisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaVal	394
Db	1141	CACCTGGGAACGAACAGAGGATCCATCTGAAGGACATTCAACAGATCGCAATGACGTT	1200
QY	395	LeuArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal	409
Db	1201	CTGCGGAGGGAACAACCGAAGTACTTTGGCATCTTCAAGGTG	1245
RESULT 5	AF521176	2159 bp mRNA linear	INV 14-AUG-2002
LOCUS	Drosophila melanogaster eiger mRNA, complete cds.		
DEFINITION	AF521176		
ACCESSION	AF521176		
VERSION	AF521176.1	GI:21717645	
KEYWORDS			
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
REFERENCE	1	(bases 1 to 2159)	
AUTHORS	Moreno, E., Yan, M. and Basler, K.		
TITLE	Evolution of TNF Signaling Mechanisms. JNK-Dependent Apoptosis Triggered by Eiger, the Drosophila Homolog of the TNF Superfamily		
JOURNAL	Curr. Biol. 12 (14), 1263-1268 (2002)		
PUBMED	12176339		
REFERENCE	2	(bases 1 to 2159)	
AUTHORS	Moreno, E., Yan, M. and Basler, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-JUN-2002) Institut Molecular Biology, Uni Zurich, Winterthurerstrasse 190, Zurich, Z 8057, Switzerland		
FEATURES	Location/Qualifiers		
source	1. 2159		
	/organism="Drosophila melanogaster"		
	/mol_type="mRNA"		
	/db_xref="taxon:7227"		
CDS	641..1888		
	/codon_start=1		
	/product="eiger"		
	/protein_id="AAW76710.1"		
	/db_xref="GI:21717646"		
	/translation="MTAETLKPFITPTSANDDGFPAKATSTATAQRRTQLIPLVLGF IGLVVAIILALITWOTTRVSHLDKELKSLRVVDNLQORLGINYLDEDFEFOKEVEN ALIDYPKVDGLTDEEDDDGLDSDIADDEDVSYSDVDVGADYEDYDMLNKL NAHTGTTPTSETTEAGEGETDTSASSNDNDNVDDFTSYNAHKKQERKSRIADVRN EEONIGNHTLEKSNESATSKESPAPLHRRMRSHRHLVVRKGSLLSARSEDS RPAAHFLSKRRHQSGMYGHGDMYIGNDNERNYSOGHFOTRDGLVLTNTGLYVYA QICNNSHDQNGFIVFGDTPFLQCLNTVPTNPHKVHTCHTSLIHLERNERIHLKD IHNDRNAVLREGNRSYFGIFKY"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,22e-141	Length:	2159
Score:	2149.00	Matches:	409
Percent Similarity:	98.55%	Conservative:	0
Best Local Similarity:	98.55%	Mismatches:	0
Query Match:	99.40%	Indels:	6
DB:	3	Gaps:	1

US-09-813-329-6 (1-409) x AF521176 (1-2159)			
QY	1	MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspAspGlyPhe	20
Db	641	ATGACTGCCGAGACCTTCAAGCCGTTTATAACGCCAAACGAGTGCCCAACGATGATGGTTTT	700
QY	21	ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu	40
Db	701	CCGGCCAAAGCGACAGCAGCGCGACCCGCCAGCAGCACCCCGCAGCTGATCCCCCTG	760
QY	41	ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTrrGln	60
Db	761	GTTTTGGGGTTTCATCGGTCTGGGGCTGGTGGTTCGCAATTCGCACTAACCACTCTGGCAG	820
QY	61	ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn	80
Db	821	ACAAACGGTGTATCGCATCTCGACAAAGGAGCTGAAGAGCCTGAAGCGAGTCTCGATAT	880
QY	81	LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr	100
Db	881	CTCCAGCAGCGCTTTGGGCATAAACTATCTGGACGAGTTCGACGAGTTCCAAAGAGGATAC	940
QY	101	GluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluGluAspAsp	120
Db	941	GAGAAATGCCCTCATCGACTATCCAAAAAGTGGATGGCTCTACGGATGAGGAGGACGAC	1000
QY	121	AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer	140
Db	1001	GACGATGGCGATGGTCTGGATTCCATTGCGGACGACGAGGACGACGAGCTTAGCTATAGC	1060
QY	141	SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn	160
Db	1061	TCTGTGATGATGTTGGCGCAGACTACGAGGACTACCGGATATGTTAAATPAACTCAAC	1120
QY	161	AsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyThr	180
Db	1121	AATGCACATACCGGCACACGCCCCACATCTGAGACCATCTGTAGGGCGAGGGCGAGACG	1180
QY	181	AspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyrAsn	200
Db	1181	GACAGTGCATCTCAGCCTCAAATGATGCAATGTGTTCGATGACTTTACGAGCTACAAT	1240
QY	201	AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu	220
Db	1241	GCCCACAAAAAGAACGAGGAGAGAAAATCTCGTCCGATTCGCCGATGTACCAATGAGGAG	1300
QY	221	GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlnLysSerSerAsnGluAlaThrSer	240
Db	1301	CAGAAATATTCAAGGAAATCACACAGAGCTTCAGGAAAAGTCATCCCATGAGGCAACTTCC	1360
QY	241	LysGluSerProAlaProLeuHisHisArgArgMetHisSerArgHisHisArgHisLeu	260
Db	1361	AAAGAGACCTTGACCACTTCACCACCGTCGCAGAAATGTCATTCGCCCATCGCCACCTC	1420
QY	261	LeuValArgLys-----AlaArgSerGluAspSerArgProAlaAla	274
Db	1421	CTAGTCCGCAAGGTGAATCTCTTCTTTCAGCCAGATCCGAGGACTCGAGGCCAGCAGCC	1480
QY	275	HisPheHisLeuSerSerArgArgHisGlnGlySerMetGlyTyrHisGlyAspMet	294
Db	1481	CATTTCCACCTGAGCAGCAGCGCGCTCACCAAGGAAGTATGGGCTTACCATGGAGATATG	1540
QY	295	TyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGly	314
Db	1541	TACATAGGAAATGATAACGAGAGAAACTCTTTATCAGGGACACTTTTCAAACCGCGCATGC	1600
QY	315	ValLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsn	334
Db	1601	GTCCTGACCGTGACCAATACAGCCCTATATTACGTATACGCCCAGATATGCTACAAACAC	1660
QY	335	SerHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeu	354
Db	1661	TCGCACGACCAAGACGATTTATCGTCTTTCAAGGAGACACTTCATCTCTGCACTGCTCTG	1720

QY 355 AnThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuile 374  
 |||||  
 Db 1721 AACCGGTGCCCAACATGCCACATAGGTGCACACCTGCCACACGAGTGTCTGATC 1780  
 |||||  
 QY 375 HisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaVal 394  
 |||||  
 Db 1781 CACCTGGAAACGAAACGAGAGGATCCATCTGAAGGACATTCACACGATCGCAATGCAGTT 1840  
 |||||  
 QY 395 LeuArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409  
 |||||  
 Db 1841 CTGGGGAGGGAACACCGAGCTACTTTGGCATCTTCAAGGTG 1885  
 |||||

## RESULT 6

AY119233 1221 bp mRNA linear INV 16-JUN-2002  
 LOCUS Drosophila melanogaster SD18286 full insert cDNA.  
 DEFINITION AY119233  
 ACCESSION AY119233  
 VERSION AY119233.1 GI:21430829  
 KEYWORDS FLI CDNA.

## SOURCE

ORGANISM Drosophila melanogaster (fruit fly)  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 1221)  
 Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,  
 Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,  
 George,R., Gonzalez,M., Guarin,H., Krommiller,B., Li,P., Liao,G.,  
 Miranda,A., Mungall,C.J., Nunoo,J., Pacleib,J., Paragas,V., Park,S.,  
 Patel,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.  
 and Celniker,S.

## REFERENCE

AUTHORS Direct Submission  
 Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,  
 Lawrence Berkeley National Laboratory, One Cyclotron Road,  
 Berkeley, CA 94720, USA  
 Sequence submitted by:  
 Berkeley Drosophila Genome Project  
 Lawrence Berkeley National Laboratory  
 Berkeley, CA 94720

## TITLE

Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,  
 Lawrence Berkeley National Laboratory, One Cyclotron Road,  
 Berkeley, CA 94720, USA

## COMMENT

This clone was sequenced as part of a high-throughput process to  
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
 Science 2000). The sequence has been subjected to integrity checks  
 for sequence accuracy, presence of a polyA tail and contiguity  
 within 100 kb in the genome. Thus we believe the sequence to  
 reflect accurately this particular cDNA clone. However, there are  
 artifacts associated with the generation of cDNA clones that may  
 have not been detected in our initial analyses such as internal  
 priming, priming from contaminating genomic DNA, retained introns  
 due to reverse transcription of unspliced precursor RNAs, and  
 reverse transcriptase errors that result in single base changes.  
 For further information about this sequence, including its location  
 and relationship to other sequences, please visit our Web site  
 (http://fruitfly.berkeley.edu) or send email to  
 cdna@fruitfly.berkeley.edu.  
 Location/Qualifiers

## FEATURES

source  
 1..1221  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"

## gene

1..1221  
 /gene="CG12919"

## CDS

170..955  
 /db\_xref="FLYBASE:FBgn0033483"  
 /gene="CG12919"  
 /note="Longest ORF"  
 /codon\_start=1  
 /product="SD18286p"  
 /protein\_id="AAM51093.1"  
 /db\_xref="GI:21430830"  
 /db\_xref="FLYBASE:FBgn0033483"  
 /translation="MLNKNNAHTGTTPTSETTAAGGETDSASSASNDNDVDFDFTS  
 YNAHKKQERKRSIADVRNEEQINQNHTELOEKSSNEATSKEAPLHRRMRHSR"

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,78e-105 Length: 1221  
 Score: 1626.00 Matches: 309  
 Percent Similarity: 97.78% Conservative: 0  
 Best Local Similarity: 97.78% Mismatches: 1  
 Query Match: 75.21% Indels: 7  
 DB: 3 Gaps: 1  
 US-09-813-329-6 (1-409) x AY119233 (1-1221)  
 QY 100 TyClnAsnAlaLeuIleAAspTyrProLysLysValAspGlyLeuThrAspGluGluAsp 119  
 |||||  
 Db 6 TACGAGAAATGCCCTCATCGACTATCCCAA-AAGGTGGATGGCCTCAGCGATGAGGAGGAC 64  
 |||||  
 QY 120 AspAspAspGlyAAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyr 139  
 |||||  
 Db 65 GACGACCATGGCGATGGTCTGGATTCCATTGGCGACGACGAGGACGACGAGCTTAGCTAT 124  
 |||||  
 QY 140 SerSerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeu 159  
 |||||  
 Db 125 AGCTCTGTGGATGATGTTGGCGAGACTACGAGGACTACACCGATATGTTAAATAAATC 184  
 |||||  
 QY 160 AsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGlu 179  
 |||||  
 Db 185 AACAAATCCATATCCCGCACCCACGCTCTGAGACCACTGCTGAGGGCGAGGGCGAG 244  
 |||||  
 QY 180 ThrAspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyr 199  
 |||||  
 Db 245 ACGGACAGTGCATCCTCAGGCTCAATGATGACNATGTGTTCCGATGACTTTTACCAGGCTAC 304  
 |||||  
 QY 200 AsnAlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGlu 219  
 |||||  
 Db 305 AATGCCCAAAAAGAGAGGAGAGAAATCTCGCTCGATTGCCGATGTACGCAATGAG 364  
 |||||  
 QY 220 GluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnLysSerSerAsnGluAlaThr 239  
 |||||  
 Db 365 GAGCAGAAATATTCAGGAAATCACACAGAGCTTCAGGAAAGTCATCCAATGAGGCAACT 424  
 |||||  
 QY 240 SerLysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHis 259  
 |||||  
 Db 425 TCCAAAGAGAGCCCTGCACCACTTCACACCGTCGCAGAAATGCATTCCTCCGCCATCGCAC 484  
 |||||  
 QY 260 LeuLeuValArgLys-----AlaArgSerGluAspSerArgProAla 273  
 |||||  
 Db 485 CTCTAGTCGCAAAAGGTGAATCTCTTTTTCAGCCAGATCCGAGGACTCGAGGCCAGCA 544  
 |||||  
 QY 274 AlaHisPheHisLeuSerSerArgArgArgHisGlnGlySerMetGlyTyrHisGlyAsp 293  
 |||||  
 Db 545 GCCCATTTTCACCTGAGCAGCAGCGCGCGTCACCAAGAAAGTATGGGCTACCATGGAGAT 604  
 |||||  
 QY 294 MetTyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAsp 313  
 |||||  
 Db 605 ATGTACATAGAAATGNTAACGAGAGAACTCTTTATCAGGGACACTTTTCAACCGCGGAT 664  
 |||||  
 QY 314 GlyValLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsn 333  
 |||||  
 Db 665 GGGCTCTTGACGGTGCACCAATACAGGCTATATTACGTATACGCCACAGATATGCTACAAC 724  
 |||||  
 QY 334 AsnSerHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCys 353  
 |||||  
 Db 725 AACTCGCACACACAGAACCGGATTTATCGTCTTTTCAAGGAGACACTCCCATTTCTGTCAGTGC 784  
 |||||  
 QY 354 LeuAsnThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeu 373  
 |||||  
 Db 785 TTGAACACGGTGGCCCAACATGCCACATPAAGGTGCACACCTGCGCACACAGTGGTCTG 844  
 |||||  
 QY 374 IleHisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAla 393  
 |||||

HRLLVRKGBESLISARSEDSPAAHFHLSRRRRHQSGMYGHDWYIIGNDNRNSYQGH  
 FQTRDGLVTVNTGLYVYVYAIQVNNSHDQNGFTVFOQDTPFLQCLNTVPTNMPHKVH  
 TCHTSGLIHLERNERIHLKDIHNDRNAVLREGNRSYFGIFKV

Db 845 ATCCACCTGGAACGAAACGAGGATCCATCTGAAGGACATTCACAAAGATCGCAATGCA 904  
 Qy 394 ValLeuArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409  
 Db 905 GTTCTCGGGAGGAAACAAACGGAAGTACTTTGGCATCTTCAAGGTG 952  
 RESULT 7  
 AC012854  
 LOCUS Drosophila melanogaster, 55359 bp DNA linear HTG 03-NOV-1999  
 DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
 ACCESSION AC012854  
 VERSION AC012854.1 GI:6223467  
 KEYWORDS HTG; HTGS, PHASE2.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 55359)  
 AUTHORS Adams, M. and Venter, J.C.  
 DIRECT SUBMISSION  
 TITLE Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,  
 JOURNAL Rockville, MD, USA  
 COMMENT This sequence was identified as CDM:10210834 by the submitter.  
 For further information on this sequence you may e-mail to  
 fly@celera.com.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 FEATURES  
 source  
 1..55359  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6.2e-95 Length: 55359  
 Score: 1508.50 Matches: 351  
 Percent Similarity: 43.01% Conservative: 0  
 Best Local Similarity: 43.01% Mismatches: 5  
 Query Match: 69.77% Indels: 463  
 DB: 2 Gaps: 5  
 US-09-813-329-6 (1-409) x AC012854 (1-55359)  
 Qy 54 LeuAlaLeuThrIleTyrGlnThrThrArgValSerHisLeuAspLysGluLeuLysSer 73  
 Db 31678 CTTATCTTGCAAGATCTGGCAGACAAACGCGTGTATCGCATCTGGCAAGGAGCTGAAGAGC 31737  
 Qy 74 LeuLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPhe 93  
 Db 31738 CTGAAGCGAGTCGTGATATCTCCACAGCGTTGGGCATAAATCTATCTGGACGAGTTC 31797  
 Qy 94 AspGluPheGlnLysGlu----- 99  
 Db 31798 GACGAGTTCCAAAAGGAGGTGAGCAATTGCGTAAATCACTTTCGAGCATCAATATTGTA 31857  
 Qy 99 ----- 99  
 Db 31858 CTCGGCGATGGGAGGGGCTGTAACATTTGCCCGGGCAATACAGATACCTCGAATTG 31917  
 Qy 99 ----- 99  
 Db 31918 TTTACAAAGTCGGAATGATTGTTATTCCTTGAGGCTGGTACTTGGATTTTAAATGAGGG 31977  
 Qy 99 ----- 99  
 Db 31978 TATATCGAACTTTTGTATTGAGTATCTATATGTTTAGTATTTCTTTTGGCCTAAAGGCT 32037  
 Qy 99 ----- 99

Db 32038 AATTGCAATTTATGATGGCAAACTCTGGCTAGGGTATCTCACAGTCGGTAAGCCATGCG 32097  
 Qy 99 ----- 99  
 Db 32098 CCTGTTTTTGAAGAAAGATGCCTCAAAAATAATGACCAGAGGTAATATATGTTGTTTGTGT 32157  
 Qy 99 ----- 99  
 Db 32158 GGGACATACAAAATATTTATAGAATTATGTGCTAAATAATAATTTTCGAGCGCTGTTTCAT 32217  
 Qy 99 ----- 99  
 Db 32218 CTGTTATTGTTTCAATTTTTTTTTTTGTTAGCTGCTGGCCAAAGGCACATAGAAAATTTGTG 32277  
 Qy 99 ----- 99  
 Db 32278 ATAGATTATTGAATTATAATGTCCTTTCCGCAAGGAACACTCATTTCAAATATGAGG 32337  
 Qy 99 ----- 99  
 Db 32338 CGCCAGCCAGTTACGTAATTTTATGAGATACAAATCCCAAGTGACAAAAGGTAAAAAGCC 32397  
 Qy 99 ----- 99  
 Db 32398 CCATTAATTCACAAAAATACTAGGCTCTATCAGTAGATATAGTCGTTCAAGCGCAGCGA 32457  
 Qy 99 ----- 99  
 Db 32458 TCGCTTTTGGCTGTATGTGACGATTGATCGAATCTCATCAAAATGCGCTAATGCGCAGAGA 32517  
 Qy 99 ----- 99  
 Db 32518 TGGTTTTAAGATTTTAAGATTTTATTATATTTTGTATGCTGGGCACTCTGTGGTGCCAAAT 32577  
 Qy 99 ----- 99  
 Db 32578 CTCTTGAGAGTTTCCTTCGCTTTTAAATATATGATATGATGTACACAAGTATATGTATAGC 32637  
 Qy 99 ----- 99  
 Db 32638 AGTATCATGTTTTTGAGCGAGGCGGCATCAATTAGGATGATGACGAAAACAGCCTGTGCTG 32697  
 Qy 99 ----- 99  
 Db 32698 CGGGCGTGTCTTAAAAATGAACGTCCCAATGGATGTGAGATATGTGCCAGGACTGTGC 32757  
 Qy 99 ----- 99  
 Db 32758 CAGCTGCAGTCACGTACTTGGCCAAAGTCTTCTGACCCAAAATGCTGTATGCCATCCCTTC 32817  
 Qy 100 -----TyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGly 117  
 Db 32818 GTTGCAATACGAGAAATGCCCTCATCGACTATCCAAAAAGGTGATGGCCTCACGGATGA 32877  
 Qy 117 uGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspAspVa 137  
 Db 32878 GGAGGACGACGACGATGCGGATGGTCTGGATTCCATTTCGCGACGACGAGGACGACGCT 32937  
 Qy 137 lSerTyrSerSerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAs 157  
 Db 32938 TAGCTATAGCTCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAA 32997  
 Qy 157 nLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGly 177  
 Db 32998 TAACTCAACAAATGCACATACCGCACCCACCCACACTCTGAGACCACTGCTGAGGGCGGA 33057  
 Qy 177 uGlyGluThrAspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheTh 197  
 Db 33058 GGGCGAGACGACAGTGCATCTCTCAGCTCAATGATGACAATGTGTTCGATGACTTTAC 33117  
 Qy 197 rSerTyrAsnAlaHisLysLysLysGlnGlu----- 207  
 Db 33118 CAGCTACAATCCCAAAAAAGAGACGAGGA-GAGGTGAGCAGCATGAAATTTTTCAGGGA 33176



QY 74 LeuLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrIleuAspGluPhe 93  
Db 131484 CTGACGCGAGTCTGCATATCTCCAGCAGCGGTTTGGGCATAACTATCTGGACGAGTTC 131425  
QY 94 AspGluPheGlnLysGlu----- 99  
Db 131424 GACGAGTTCCAAAGGAGGTGAGCAATTCGGTAATCACTTCGAGCATCAATATTGTGA 131365  
QY 99 ----- 99  
Db 131364 CTCGGCGATGGGCGCAGGGCTGTAAACATTTGCCCGGCAAAATACAGATACCTCGAATTG 131305  
QY 99 ----- 99  
Db 131304 TTTACAAAGTCGGAATGATTGTATTCTCTGAGGCTGTACTTGAGTTTAAATGAGGG 131245  
QY 99 ----- 99  
Db 131244 TATATCGAACTTTTGTATTGAGTATCTATTATGTATTGTCTTTTTCCTTAAGGGT 131185  
QY 99 ----- 99  
Db 131184 AATTTCGATTTTATGATGCCAAACTCTGGCTAGGGTATCTCACAGTCGGTAAGCCATGCG 131125  
QY 99 ----- 99  
Db 131124 CCTGTTTTGAAAAGAAATGCCTCAAAAATGACCAGACGTAATATGTGTTATTGTGT 131065  
QY 99 ----- 99  
Db 131064 GGGACATACAAAATATTATAGAAATTATGTGCTAAATATAAATTTCGACGCTGTTTCAT 131005  
QY 99 ----- 99  
Db 131004 CTGTTATTGTTTCATTTTGTGTTAGCTGTTGGCCAAAGGCACATAGAAATTTGTG 130945  
QY 99 ----- 99  
Db 130944 ATAGATTATTGAATTATAATGTCTCTTCGCAAGGAACACTCAITTCAAATATGAGG 130885  
QY 99 ----- 99  
Db 130884 CGCCAGCCAGTTCAGTAAATTATGAGATACAAATCCCAAGTGACAAAAGGTAAAGCC 130825  
QY 99 ----- 99  
Db 130824 CCAATAATTCACAAAATACTAGGCTCTATCAGTAGATATAGTCGTTCAAGCGGAGCGGA 130765  
QY 99 ----- 99  
Db 130764 TCGCTTTTGGCTGTATGTGACGATTGATCGAATCTCATCAATGGCCTAATGGCCAGAGA 130705  
QY 99 ----- 99  
Db 130704 TGGTTTTAAGATTTAAGATTATTATATTTTGAATGGCTGGGCATCTGTGTGGTCCAAAT 130645  
QY 99 ----- 99  
Db 130644 CTCTTGAGAGTTTCTCTTCGCTTTTAAATATATGTATGATGTACACAAGTATATGTATAGC 130585  
QY 99 ----- 99  
Db 130584 AGTATCATGTTTTTGAAGCGGCGCCATCAATTAGGATGATGACGAAACAGCCTGTGTCTG 130525  
QY 99 ----- 99  
Db 130524 CGGCGGTGTCTTAAAAATGAACGTCCCAATCGATGTGAGATATGTGCCAGGACTGTGC 130465  
QY 99 ----- 99  
Db 130464 CAGTCGAGTCAGGTACTTTGGCCAAAGTCTTCTGACCCAAAATGCTGTATGCCCATCCCTTC 130405

QY 100 -----TyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGly 117  
Db 130404 GTTTCAGTACGAGAAATGCCCTCATCGACTATCCAAAAAGGTGATGGCTCTACCGATGA 130345  
QY 117 uGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspVa 137  
Db 130344 GGAGGACGACGAGATGGCGATGGTCTGGAATTCATTTGGCGACGACGAGACGACGCGT 130285  
QY 137 lSerTyrSerSerValAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAs 157  
Db 130284 TAGCTATAGTCTCTGTGATGATGTTGGCGCAGACTACGAGGACTACCCGATATGTTAAA 130225  
QY 157 nLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyG 177  
Db 130224 TAAACTCAACAATGACATACCGGCACACCGCCACCATCTGAGACCATGCTGAGGCGCA 130165  
QY 177 uGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheTh 197  
Db 130164 GGGCGAGACGGACAGTGCATCTCAGCCTCAAAATGATGACAAATGTGTTGATGACTTTAC 130105  
QY 197 rSerTyrAsnAlaHisLysLysGlnGlu----- 207  
Db 130104 CAGCTACAATGCCACAAAAGAACGACGA - GAGGTGAGCACGGATGAAATTTTCAGGA 130046  
QY 208 -----ArgLysSerArgSerIleAlaAspVa 216  
Db 130045 TTTATCGTATTAAAGCTCTGCTTTTCTCTTCACAGAAAATCTCGCTCGATTGCCGATGT 129986  
QY 216 lArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlyLysSerSerAs 236  
Db 129985 ACGCAATGAGGAGCAGATATTTCAAGGAATTCACACAGACTTCAGGAAAAGTCTATCAA 129926  
QY 236 nGluAlaThrSerLysGlu----- 242  
Db 129925 TGAGGCAACTTCCAAAGAGAGGTAAGTAGAACGTTCCGTTGAAAAACAGCAAGAAAAACC 129866  
QY 242 ----- 242  
Db 129865 ATTAAATGAAGTTATATGCTATGGGTTAAAGATAGATAGTTAACTGCGTAACTGAACACA 129806  
QY 243 -----SerProAlaProLeuHisHisArgArgArgMeth 254  
Db 129805 TATGTATACATTTTCTTTTACTTGAAGCCCTGCGACCACTTCACCACCGTCCGAGATGC 129746  
QY 254 lSerArgHisArgHisLeuLeuValArgLysAla----- 265  
Db 129745 ATTCCCGCATCGCCACCTCTCTAGTCCGCAAAAGG - TGTGTAGTGTAGAAAAACGCAACAA 129687  
QY 265 ----- 265  
Db 129686 TTGGCAAAAATCGAATACCAATCTCTATTCCGACCTCTCTTAAATTTGCGAGGTGAATCTC 129627  
QY 266 -----ArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgA 282  
Db 129626 TTTCTTCAGCCAGATCCGAGGACTCGAGGCCAGCAGCCCACTTCCACTTGAGCAGCAGGC 129567  
QY 282 rGArgHisGlnGly----- 286  
Db 129566 GCGGTCCACCAAGGAAGTATGGTAAAGAGATGTTCCGCGCTCAAGCGGTGACATACAAGT 129507  
QY 287 -----SerMetGlyTyrHisGlyAspMetT 295  
Db 129506 ACTACAGCAACCTAATGAAGTCACTTCTTGAATTCTCA - GGTACCATGGAGATATGT 129448  
QY 295 yrlleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyV 315  
Db 129447 ACATAGAAAATGATAACGAGAGAAACTCTTATCAGGGACACTTTCAAACCGCGATGGCG 129388  
QY 315 alLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsn 335  
Db 129387 TCTTGACGGTGACCAATACAGGCTTATTTACGTATACGCCCCAGATATGTCACAAACT 129328  
QY 335 erHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuA 355

```
Db 129327 CGCAGCAGCAGGATTTATCGTCTTTCAGGAGACACTCCATTCCTGCGAGTGTGA 129268
Qy 355 anThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleH 375
Db 129267 ACACGGTGCCCAACACATGCAATAGGTGCGACACTGCGCACAGAGTGGTCTGATCC 129208
Qy 375 isLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValL 395
Db 129207 ACCTGGACGAACAGAGAGGATCCATCTGAAGGACATTCACACGATCGCAATGAGTTC 129148
Qy 395 euArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409
Db 129147 TCGCGGAGGGAACAACACGAGTACTTTGCGATCTTCAAGGTG 129104

RESULT 9
AC007414
LOCUS AC007414 172904 bp DNA linear INV 06-SEP-2001
DEFINITION Drosophila melanogaster, chromosome 2R, region 46C-46D, BAC clone
BACR06J01, complete sequence.
ACCESSION AC007414
VERSION AC007414.6 GI:15451491
KEYWORDS HTG.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 172904)
Celnikier,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Ananidis,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferreira,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwan,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Swirskas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 2R, region 46C-46D
Unpublished
2 (bases 1 to 172904)
Celnikier,S.E., Agbavani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
Swirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
Submitted (28-APR-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Sep 6, 2001 this sequence version replaced gi:13324748.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdgp@fruitfly.berkeley.edu.
Location/Qualifiers
1..172904
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

FEATURES
source
1..172904
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
```

```
/db_xref="taxon:7227"
/chromosome="2R"
/map="46C-46D"
/clone="BACR06J01 (D601)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6)"

ORIGIN
Alignment Scores:
Pred. No.: 2.27e-94 Length: 172904
Score: 1508.50 Matches: 351
Percent Similarity: 43.01% Conservative: 0
Best Local Similarity: 43.01% Mismatches: 5
Query Match: 69.77% Indels: 463
DB: 3 Gaps: 5

US-09-813-329-6 (1-409) x AC007414 (1-172904)
Qy 54 LeuAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAepLysGluLeuLysSer 73
Db 136338 CTTATCTTGCAGATCTGGCAGACACACGCGTGTATCGCATCTGGACAGGAGCTCAAGAGC 136397
Qy 74 LeuLysArgValValAepAsnLeuGlnArgLeuGlyIleAenTyrLeuAepGluPhe 93
Db 136398 CTGAAGCGAGTCGTCGATAATCTCCAGCAGCGTTTGGGCATAAACTATCTGGACGAGTTC 136457
Qy 94 AspGluPheGlnLysGlu----- 99
Db 136458 GACGAGTTCCTCAAAAGGAGGTGAGCAATTCGCTAAATCACTTTCGAGCATCAATATTGTA 136517
Qy 99 ----- 99
Db 136518 CTGGCGATGGGCGAGGGGCTGTAAACATTTGCCCGGCAAAATACAGATACCTCGAATTG 136577
Qy 99 ----- 99
Db 136578 TTTACAAAGTCGGAATGATTGTTATTCCTTGAGGCTGTACTTGAGTTTTTAATGAGGG 136637
Qy 99 ----- 99
Db 136638 TATATCGAACTTTTGTATTATGAGTATCTATTATGTTAGTTTCTTTTCCTTAAAGGGT 136697
Qy 99 ----- 99
Db 136698 AATTGCAATTTATATGCGCAAACTCTGGCTAGGGTATCTCACAGTCGGTAAGCCATGCG 136757
Qy 99 ----- 99
Db 136758 CCTGTTTTTGAAAAAGAAATGCCTCAAAAAAATGACACAGAGCTAATATGTGTTATTGTGT 136817
Qy 99 ----- 99
Db 136818 GGGACATACAAAAATATTTATAGAAATTATGTGCTAAATAATAATTTTCGACGCTGTTTCAT 136877
Qy 99 ----- 99
Db 136878 CTGTTATTGTTTCATTTTTTTTTTTTGTAGCTGCTTGCCCAAGGCACATAGAAAAATTGTG 136937
Qy 99 ----- 99
Db 136938 ATAGATTATTGGAATTATTAATGTCCTCTTCCGCAAGGAACACTCATTTCAAATATGAGG 136997
Qy 99 ----- 99
Db 136998 CGCCAGCCAGTTCACGTAATTTTATGAGATACAAATCCCCAAGTGACAAAAAGGTAAGGCC 137057
Qy 99 ----- 99
Db 137058 CCAATAATTCAAAAAAATACTAGGCTCTATCATGATAGATATAGTCGTTCAAGGCGAGCGA 137117
Qy 99 ----- 99
```

Db	137118	TCGCTTTTGGCTGATGTGACGATTGATCGAATCTCATCAAAATGCCTAATGGCCAGAGA	137177
Qy	99	-----	99
Db	137178	TGGTTTTAAAGATTTTAAAGATTTATTTATATTTTGTGGCTGGGCATCTGTGTGCCAATT	137237
Qy	99	-----	99
Db	137238	CTCTTGAGAGTTTCCTTCGCTTTTAAATAATATGTATGTATACAAAGTATATGTATACG	137297
Qy	99	-----	99
Db	137298	AGTATCATGTTTTTGAGCGAGCGCCATCAATTAGGATGATGACGAACACGCTGTGCTG	137357
Qy	99	-----	99
Db	137358	CGGGCGTGTGCTAAAAAAATGAACGTCCCAAAATGGATGTGAGATATGTGCCAGGACTGTGC	137417
Qy	99	-----	99
Db	137418	CAGCTGCAGTCACGTACTTGGCCAACTCTTGACCCAAAATGCTGATGCCCATCCCTTC	137477
Qy	100	-----TyrGluAsnAlaLeuLeuAspTyr:ProlylsLysValAspGlyLeuThrAspG1	117
Db	137478	GTTGCAGTACGAGAATGCCCTCATCGACTATCCAAAAAAGGTGGATGSCCTCACGGATGA	137537
Qy	117	uGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspVa	137
Db	137538	GGAGGACGACGAGTGGCGATGGTCTCGATTCATTCGGACACGAGGACGACGACGT	137597
Qy	137	IserTyrSerValAspAspValGlyValaAspTyrGluAspTyrThrAspMetLeuAs	157
Db	137598	TAGCTATAGTCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAA	137657
Qy	157	nLysLeuAsnAsnAlaHisThrGlyThrProThrSerGluThrThrAlaGluGlyG1	177
Db	137658	TAAACTCAACATGCACATACCGGCACACGCCACACATCTGAGACCATCTGCTGAGGCGCA	137717
Qy	177	uGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspValPheAspAspPheTh	197
Db	137718	GGCGGAGACGACAGTGCATCTCAGCTCAAAATGATGACAAATGTTTCGATGACITTC	137777
Qy	197	rSerTyrAsnAlaHisLysLysLysGlnGlu-----	207
Db	137778	CAGCTACAAATGCCACAAAGAGACGGA-GAGGTGAGCACGGATGAAATTTTCAGGGA	137836
Qy	208	-----ArgLysSerArgSerIleAlaAspVa	216
Db	137837	TTTATCGTATTAAGCCTCTGCTTTTCTCTTTCACAGAAAATCTGCTCGATTGCCGATGT	137896
Qy	216	lArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlyLysSerSerAs	236
Db	137897	ACGCAATGAGGAGCAGAAATATTCAGGGAATATCACAGAGCTTCAGGAAAAGTCATCCAA	137956
Qy	236	nGluAlaThrSerLysGlu-----	242
Db	137957	TGAGGCAACTTCCAAAGAGAGGTAAGTAGAACGTTCCGTTGAAAAACAGGCAAGAAAACC	138016
Qy	242	-----	242
Db	138017	ATTAATGAAGTTATATCTATGGGTTAAAGATAGATAGTTAACTGCGTAACCTGAAACACA	138076
Qy	243	-----SerProAlaProLeuHisHisArgArgArgMeth	254
Db	138077	TATGTATACATTTTCTTTTACTTGAAGCCCTGCACACCTTCACACCGCTCGCAGATGC	138136
Qy	254	iSerArgHisArgHisLeuValArgLysAla-----	265
Db	138137	ATTCCCGCCATCGCCACTCTAGTCGCAAGG-TGTGTAGTGTAGAAAAGCGAAACAA	138195
Qy	265	-----	265
Db	138196	TTGCGAAAAATCGAATACCAAAATCTCTATTTCGCACCTCTGTTAAATTGCGAGGTGATCTC	138255
Qy	266	-----ArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgA	282
Db	138256	TTCTTTTCAGCCAGATCCGAGGACTCGAGGCCAGCAGCCCATTTCCACTTTGAGCAGCAGGC	138315
Qy	282	rgArgHisGlnGly-----	286
Db	138316	GGCGTCAACCAAGGAAGATATGGTTAAGGAGATGCTTCCGGGTCAAGCGTGTACATACAAGT	138375
Qy	287	-----SerMetGlyTyrHisGlyAspMetT	295
Db	138376	ACTACAGGCAACCTAATGAAGTCACCTCTTGAATTTCTCA-GGCTACCATGGAGATATGT	138434
Qy	295	YrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyV	315
Db	138435	ACATAGGAATATGATACGAGAGAAACTTTATCAGGGACACTTTCAAACCGCGGATGGCG	138494
Qy	315	alleuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnS	335
Db	138495	TCTTGACGGTGACCAATACAGGCCTATATTACGTATACGCCAGATATGCTACAACAACT	138554
Qy	335	erHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuA	355
Db	138555	CGCACGACGAGAAACGATTTATCGTCTTTCAGGAGACACTCCATCTCTGCGAGTCTGTA	138614
Qy	355	snThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleH	375
Db	138615	ACACGGTGGCCCAACATGCCACATAAGGTGCACACCTGCCACACGAGTGGTGTGATCC	138674
Qy	375	isLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaVal	395
Db	138675	ACCTGGACGAAACGAGAGGATCCATCTGAGGACATTCACAAACGATCGCAATGAGTTC	138734
Qy	395	euArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal	409
Db	138735	TGCGGAGGAAACAAACCGAGCTACTTTGGCATCTTCAAGGTG	138778
RESULT 10			
AC099029/c		189620 bp	DNA linear INV 08-NOV-2001
LOCUS			
DEFINITION			Drosophila melanogaster, chromosome 2R, region 46D-47A, BAC clone
			BACR09J12, complete sequence.
AC099029			
VERSION			AC099029.1
KEYWORDS			GT:16798961
SOURCE			HTG.
ORGANISM			Drosophila melanogaster (fruit fly)
REFERENCE			Drosophila melanogaster
AUTHORS			Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
			1 (bases 1 to 189620)
			Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Bazon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferrera, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houch, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalili, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Pacle, J., Paragov, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svirska, R., Tector, C., Williams, S.M., Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.
			Sequencing of Drosophila chromosome 2R, region 46D-47A
			Unpublished
TITLE			2 (bases 1 to 189620)
JOURNAL			Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Bazon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferrera, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A.,
REFERENCE			
AUTHORS			

Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Hostin,D., Howland,T.J., Ibegwan,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

## Direct Submission

Submitted (08-NOV-2001) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US

## Sequence submitted by:

Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory, MS 64-121  
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to [bdgp@fruitfly.berkeley.edu](mailto:bdgp@fruitfly.berkeley.edu).

## Location/Qualifiers

1..189620  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/strains="y; cn bw sp"  
/db\_xref="taxon:7227"  
/chromosome="2R"  
/map="46D-47A"  
/clone="BACR09J12 (D1345)"  
/clone\_lib="RPCI-98 (Roswell Park Cancer Institute  
Drosophila melanogaster BAC library, partial EcoRI in  
pBACe3.6)"

## ORIGIN

## Alignment Scores:

Pred. No.: 2.52e-94 Length: 189620  
Score: 1508.50 Matches: 351  
Percent Similarity: 43.01% Conservative: 0  
Best Local Similarity: 43.01% Mismatches: 5  
Query Match: 69.77% Indels: 463  
DB: 3 Gaps: 5

US-09-813-329-6 (1-409) x AC099029 (1-189620)

Qy 54 LeuAlaLeuThrIleTrrGlnThrThrArGValSerHisLeuAspLysGluLeuLysSer 73  
Db 166400 CTTATCTTGCAGATCTGCAGACAAACGCGTGTATCGCATCTGGACAAGGAGCTGAAGAGC 166341  
Qy 74 LeuLysArgValValAspAsnLeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPhe 93  
Db 166340 CTGAAGCGAGTCGTGATAAATCTCAGCAGCGTTTGGCATAAACTATCTGGACGAGTTC 166281  
Qy 94 AspGluPheGlnLysGlu----- 99  
Db 166280 GACGAGTTCCAAAAGAGGCTGAGCAATTGCGTAATATCACTTTCGAGCATCAATATTTGTA 166221  
Qy 99 ----- 99  
Db 166220 CTCGGCGATGGGCGAGGGGTGTAACATTGCCCCGGGCAAAATACAGATACCTCGAATG 166161  
Qy 99 ----- 99  
Db 166160 TTTACAAAGTCGGAATGATTGTTATTCCTTGGAGCTGGTACTTGAGTTTTTAATGAGGG 166101  
Qy 99 ----- 99  
Db 166100 TATATCGAACTTTTGTATTGAGTATCTATTATGTTTAGTTTCTTTTGTGCTAAAGGGT 166041  
Qy 99 ----- 99  
Db 166040 AATTTGCATTTTATGATGCAAACTCTGGCTAGGGTATCTCACAGTCGGTAAGCCATGCG 165981

Qy 99 ----- 99  
Db 165980 CCTGTTTTGAAAAAGAATGCCTCAAAAAAATGACCAGACGTAATATGTGTTATTGTGT 165921  
Qy 99 ----- 99  
Db 165920 GGGACATACAAAAATATTTATAGAATTTATGTGCTAAAAATATAAATTTCCAGCGCTGTTTCAT 165861  
Qy 99 ----- 99  
Db 165860 CTGTTATTTGTTTCATTTTTTTTTTTGTTAGCTGCTTGCCAAAGCACATAGAAAAATTGTG 165801  
Qy 99 ----- 99  
Db 165800 ATAGATTATTGAATTATAATGTCTCTTCGCAAGGAACACACTCATTTCAAATATGAGG 165741  
Qy 99 ----- 99  
Db 165740 CGCCAGCCAGTTCACGTAATTTATAGATACAAATCCCAAGTGACAAAAGGTAAAGCC 165681  
Qy 99 ----- 99  
Db 165680 CCAATAATTCAAACAAAAATACTAGGCTCTATCAGTAGATATAGTCGTTCAAGGCGAGCGA 165621  
Qy 99 ----- 99  
Db 165620 TCGCTTTTGGCTGTATGTGACGATTGATCGAATCTCATCAATGGCCTAATGGCCAGAGA 165561  
Qy 99 ----- 99  
Db 165560 TGGTTTTAAGATTTTAAGATTTATTTATATTTTGATGGCTGGGCATCTGTGGTCCCAATT 165501  
Qy 99 ----- 99  
Db 165500 CTCTTGAGAGTTTCCTTCGCTTTTAAATATATGTATGATGTACACAAGTATATGTATACG 165441  
Qy 99 ----- 99  
Db 165440 AGTATCATGTTTTTTGAGCGAGGCGCCATCAATTAGGATGATGACGAAACAGCCTGTGCTG 165381  
Qy 99 ----- 99  
Db 165380 CGGCGGTGTGCTTAAAAAATGAACGTCGCCAAATGGATGTGAGATATGTGCCAGGACTGTGC 165321  
Qy 99 ----- 99  
Db 165320 CAGCTGCAGTCACGTACTTGGCCAGTCTTCTGACCCAAAATGCTGATGCCATCCCTTC 165261  
Qy 100 -----TyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspG1 117  
Db 165260 GTTCAGTACGAGAAATGCCCTCATCGACTATCCAAAAAAGGTGGATGGCCTCACGGATGA 165201  
Qy 117 uGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspAspVa 137  
Db 165200 GGAGGACGACGACGATCGCATGGTCTGSGATTCCATTGCCGACGACGAGGACGACGCT 165141  
Qy 137 lSerTyrSerSerValAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAs 157  
Db 165140 TAGCTATAGCTCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAA 165081  
Qy 157 nLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyG1 177  
Db 165080 TAAACTCAAAATGCATACACCGCACCCACATCTGAGACCCTGCTGAGGCGCA 165021  
Qy 177 uGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheTh 197  
Db 165020 GGGCGAGACGACAGTGCATCTCCAGCCTCAAAATGATGACAAATGTGTTCGATGACTTTAC 164961  
Qy 197 rSerTyrAsnAlaHisLysLysLysGlnGlu----- 207  
Db 164960 CAGCTACATGCCCAAAAAGAAAGACAGGA- GAGGTGAGCAGCGATGAAATTTTCAGGGA 164902  
Qy 208 -----ArgLysSerArgSerIleAlaAspVa 216



Db	164901	TTTATCGTATTAAAGCCTCTGCTTTTCTCTTACAGAAATCTCGCTCCATTGCCGATGT	164842
QY	216	lArGAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluYSerSxAs	236
Db	164841	ACGCAATGAGGACGAGCAATATTCAAGGAAATCACAGAGCTTCAGGAAAGTCATCCAA	164782
QY	236	nGluAlaThrSerLysGlu-----	242
Db	164781	TGAGGCACTTCCAAAGAGAGGTAGTAACGTTCCGTTGAAAAACAGCGCAAAAAACC	164722
QY	242	-----	242
Db	164721	ATTAATGAAGTTATATGCTATGCTGGTTAAAGATAGTAGTAACTCGGTAACTGAACACACA	164662
QY	243	-----SerProAlaProLeuHisHisArgArgMeth	254
Db	164661	TATGTATACATTTCTTTTACTTGAAGCCCTGCACCACTTCACCACTGCAGATGC	164602
QY	254	isSerArgHisArgHisLeuLeuValArgLysAla-----	265
Db	164601	ATTCCGCCATCGCCACCTCTTAGTCCGCAAGG-TGTGTAGTGTAGAAAAAGCGAAACAA	164543
QY	265	-----	265
Db	164542	TTGGCAAAAATCGAATACCAATCTCTATTTCACCTCTGTTTAATTTTCAGGTTGAATCTC	164483
QY	266	-----ArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgA	282
Db	164482	TTCTTTAGCCAGATCCGAGGACTCGAGCCAGCAGCCCATTTCCACTGTGACGACGCG	164423
QY	282	tGArgHisGlnGly-----	286
Db	164422	GGCGTCAACAGGAAGTATGGTAGGAGATGCTTCGCCGTCAACGCTGTATACATCAAGT	164363
QY	287	-----SerMetGlyTyrHisGlyAspMetT	295
Db	164362	ACTACAGCAACCTAATGAAGTCACTTCTTGAATTTCTCA-GGCTACCATGGAGATATGT	164304
QY	295	yrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyV	315
Db	164303	ACATAGGAATGTATAACGAGAACTCTTATCAGGACACTTTCAACCGCGGATGGCG	164244
QY	315	alLeuThrValThrAsnThrGlyLeuTyrValTyrAlaGlnIleCysTyrAsnAsnS	335
Db	164243	TCITGACGTTGACCAATACAGGCCTATATTACGTATACGCCAGATATGCTACACAACT	164184
QY	335	erHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuA	355
Db	164183	CGCAGCACCAACGGATTTATCGTCTTTCAAGGAGACATCCATTCTCTCAGTGTCTGA	164124
QY	355	snThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleH	375
Db	164123	ACACGTTGCCCAACATGCCATAGGTGCACACTCCACACGAGTGGTCTGATCC	164064
QY	375	isLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaVal	395
Db	164063	ACCTGGACGAAACGAGAGATCCATCTGAAGGACATTCACAACGATCGCAATGCGATTC	164004
QY	395	euArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal	409
Db	164003	TGCGGGAGGGAACACCGAGCTACTTTGGCATCTTCAAGGTG	163960
RESULT 11	AE003831		
LOCUS	AB003831	275390 bp	DNA linear INV 14-FEB-2003
DEFINITION	Drosophila melanogaster chromosome 2R, section 19 of 74 of the complete sequence		
ACCESSION	AE003831	AE002787	AE013599
VERSION	AE003831.3	GI:21645483	
KEYWORDS	Drosophila melanogaster (fruit fly)		
SOURCE	Drosophila melanogaster		
ORGANISM	Drosophila melanogaster		

# REFERENCE AUTHORS

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 275390)  
Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Ananidis,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,G.L., Abail,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Beriman,B.P., Bhandari,D., Bolshakov,S., Borkov,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotier,P., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Center,A., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrieres,S., Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kennis,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusskern,D.R., Pacle,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodagef, Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.  
The genome sequence of Drosophila melanogaster  
Science 287 (5461), 2185-2195 (2000)  
20156006  
10731132

# REFERENCE AUTHORS

2 (bases 1 to 275390)  
Celniker,S.E., Adams,M.D., Kronmiller,B., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Ananidis,P.G., Brandon,R.C., Rogers,Y., Banzon,J., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferrieres,S., Frise,B., Galle,R.F., Garg,N.S., George,R.A., Gonzales,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Shapiro,S.M., Strong,R., Svirskaas,R., Tector,C., Tyler,D., Williams,S.M., Zaveri,J.S., Smith,H.O., Venter,J.C. and Rubin,G.M.  
Sequencing of Drosophila melanogaster genome  
Unpublished  
3 (bases 1 to 275390)  
Miera,S., Crosby,M.A., Matthews,B.B., Bayraktaroglu,L., Campbell,K., Hradscky,P., Huang,Y., Kaminker,J.S., Prochnik,S.E., Smith,C.D., Tupy,J.L., Bergman,C.M., Berman,B.P., Carlson,J.W., Celniker,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Frise,E., de Grey,A.D.N.J., Harris,N.L., Kronmiller,B., Marshall,B., Millburn,G.H., Richter,J., Russo,S., Searle,S.M.J., Smith,E., Shu,S., Smutniak,F., Whitfield,E.J., Ashburner,M., Gelbart,W.M., Rubin,G.M., Mungall,C.J. and Lewis,S.E.  
Annotation of Drosophila melanogaster genome  
Unpublished  
4 (bases 1 to 275390)

# REFERENCE AUTHORS

3 (bases 1 to 275390)  
Miera,S., Crosby,M.A., Matthews,B.B., Bayraktaroglu,L., Campbell,K., Hradscky,P., Huang,Y., Kaminker,J.S., Prochnik,S.E., Smith,C.D., Tupy,J.L., Bergman,C.M., Berman,B.P., Carlson,J.W., Celniker,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Frise,E., de Grey,A.D.N.J., Harris,N.L., Kronmiller,B., Marshall,B., Millburn,G.H., Richter,J., Russo,S., Searle,S.M.J., Smith,E., Shu,S., Smutniak,F., Whitfield,E.J., Ashburner,M., Gelbart,W.M., Rubin,G.M., Mungall,C.J. and Lewis,S.E.  
Annotation of Drosophila melanogaster genome  
Unpublished  
4 (bases 1 to 275390)

AUTHORS	Adams,M.D., Celnikier,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
REFERENCE	5 (bases 1 to 275390)		
AUTHORS	FlyBase		
CONSTRM	Direct Submission		
TITLE	Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA		
JOURNAL	On Jul 1, 2002 this sequence version replaced gi:10727672. Location/Qualifiers		
COMMENT	1. .275390		
FEATURES	source		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7227"		
	/chromosome="2R"		
	/note="Genotype: y[1]; cn[1] bw[1] sp[1]; Rh6[1]"		
gene	197. .2954		
	/locus_tag="CG1671"		
	/note="last curated on Mon Feb 11 17:53:43 PST 2002"		
	/map="46B7-46B9"		
mRNA	/db_xref="FLYBASE:FBgn0033454"		
	join(197. .286,346. .2954)		
	/locus_tag="CG1671"		
	/product="CG1671-RA"		
CDS	/db_xref="FLYBASE:FBgn0033454"		
	join(261. .286,346. .2683)		
	/locus_tag="CG1671"		
	/note="CG1671 gene product"		
	/codon_start=1		
	/product="CG1671-PA"		
	/protein_id="AAF58888.1"		
	/db_xref="GI:7303841"		
	/db_xref="FLYBASE:FBgn0033454"		
	/translation="MLAGISEKSYAAEARYTNFVAGGDIAWSPNGHLYCLINGSANVQ VQDTGILNSYSSADNKPRAELDNIDVEEDTITCGLSQEHVLTVHRSGLLR LWDGTGLKWAHQKGPVIRVEFPQCRLICITSGGADATLRWDYSNNSCLGALKD PGPAWLLVHPNVLKKEIYAGGSDNTVYAWNYETKLVHKMRGHLQVTLGLSPKSA DPCNEIVTSVDKVMYVQLQOELESKQLKVLPLYDELEGVYADEQLIIVASGSK LQQVDPKSKIRDLSSGDFQISRLHCSLKLQALVTTEQNIIVDVEITLPIKHLV GYNDEIILDMCPMBGHDRLYAVATNSKHFKLYDTEHDMNCKLIVGHSDTVMSLAASQNL LISYKDCSVRLMRLOHDKDCSLQALVQAQNCHTSTIGCVAMTNGATGFASVSQDGS MKVQLVRSKEDRNSYFNRYAALSHDKVCNVAYAPNNKLIATAGQDKTAKVWLAE SNTLQGVLRGTRGVSWSPSPVQIIVLTSSSDCTLRISNFPSCIKRFPDQCTILR AEFIDHGKFIISASDGILLKWNITKNTCLQSLDEHNDRVWALVARSNRFFYTGGA DSKLIRGDVTVQTRNRAIDKQAALQEQETLHSLHAQKLKAFVIALNLDKPKAS FDIINHFVRKDEPLQRLVDQLNQRVALLOHVKAWTTNSRHSQAGNMILKYLIGD ALLDPKARFYNNVNEVLTPYVQRFKRVTELNKOLAFLEFIVKCM"		
	complement(2844. .13679)		
gene	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/note="last curated on Tue Mar 12 18:58:27 PST 2002"		
	/map="46B9-46B12"		
mRNA	/db_xref="FLYBASE:FBgn0027580"		
	complement(join(2844. .3521,3590. .4619,4684. .5341, 6647. .8156,8220. .8346,13587. .13679))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-RE"		
mRNA	/db_xref="FLYBASE:FBgn0027580"		
	complement(join(2844. .3521,3590. .4619,4684. .5341, 5819. .5866,6647. .8156,8220. .8346,12506. .12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
mRNA	/db_xref="FLYBASE:FBgn0027580"		
	complement(join(2844. .3521,3590. .4619,4684. .5341, 5819. .5866,6647. .8156,8220. .8346,12506. .12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-RB"		
mRNA	/db_xref="FLYBASE:FBgn0027580"		
	complement(join(2844. .3521,3590. .4619,4684. .5341, 5819. .5866,6647. .8156,8220. .8346,12506. .12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-RK"		
mRNA	/db_xref="FLYBASE:FBgn0027580"		
	complement(join(2844. .3521,3590. .4619,4684. .5341, 5819. .5866,6647. .8156,8220. .8544))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-RG"		
mRNA	/db_xref="FLYBASE:FBgn0027580"		
	complement(join(2844. .3521,3590. .4619,4684. .5341, 5819. .5866,6647. .8156,8220. .8544))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-RL"		
CDS	/db_xref="FLYBASE:FBgn0027580"		
	complement(join(3273. .3521,3590. .4619,4684. .5341, 6647. .8156,8220. .8318))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	Alignment Scores:		
	Pred. No.:	3,86e-94	Length: 275390
	Score:	1508.50	Matches: 351
	Percent Similarity:	43.01%	Conservative: 0
	Best Local Similarity:	43.01%	Mismatches: 5
	Query Match:	69.77%	Indels: 463
	DB:	3	Gaps: 5
	US-09-813-329-6 (1-409) x AB003831 (1-275390)		
Qy	54	LeuAlaLeuThrIleTrrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSer	73
Db	233631	CTTATCTTGAGATCTGGCAGACAACCGGTGTATCGCATCTGGACAAGGAGCTGAAGAGC	233690
Qy	74	LeuLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPhe	93
Db	233691	CTGAAGCGAGTCGTGATATATCTCCAGCAGCGTTTGGGCATAACTATCTGGACGAGTTC	233750
Qy	94	AspGluPheGlnLysGlu	99
Db	233751	GACGAGTTCCAAAAGGAGGTGAGCAATTGGTAAATCACATTCGAGCATCAATATTGTA	233810
Qy	99	-----	99
Db	233811	CTCGGCGATGGGAGGGGCTGTAAACATTTGCCCGGGCAATACAGATACCTCGAATTG	233870







```
PFMAL1P2_2
WPCOMMENT
Sequence split into 4 fragments LOCUS PFMAL1P2 Accession AL031745
Fragment Name Begin End
PFMAL1P2_0 1 110000
PFMAL1P2_1 100001 210000
PFMAL1P2_2 200001 310000
PFMAL1P2_3 300001 384550
Continuation (3 of 4) of PFMAL1P2 from base 200001 (AL031745 Plasmodium falciparum DNA f

Alignment Scores:
Pred. No.: 0.463 Length: 110000
Score: 172.50 Matches: 83
Percent Similarity: 36.73% Conservative: 61
Best Local Similarity: 21.17% Mismatches: 153
Query Match: 7.98% Indels: 95
DB: 3 Gaps: 15

US-09-813-329-6 (1-409) x PFMAL1P2_2 (1-110000)
Qy 67 LeuAspLysGluLeuLysSerLeu-----LysArgValValAspAsnLeu--- 81
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72659 CTAGATAAAATGAAGAGCAACTTTAATGATGAAGATGCTGATGAGAATATGATA 72718
Qy 82 -----GlnGlnArgLeuGly-----IleAsnTyrLeuAspGluPheAspGluPhe 96
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72719 AAATTGGATAGCAGTAATAGGAATAAAGAAATTAAGAAATGAAGAAATGATGATTT 72778
Qy 97 GlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAsp 116
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72779 GAAGAAGAAGAAAGAAATAGATGTTGAC-----GATGTTGATGACGTTGACGAT 72829
Qy 117 GluGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspAsp 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72830 GTTGATGACGTTGACGATGTTGACGATGTTGACGATGTTGACGATGATGAAGATGATGAA 72889
Qy 137 ValSerTyrSerSerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeu 156
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72890 GATGATGAGGATGACGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAA 72949
Qy 157 AsnLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGly 176
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72950 GATGAAGATGATGAT-----GATGATGAAGATGAAGATGATGATGATGATGATGATGAC 72988
Qy 177 GluGlyGluThrAspSerAlaSerAsnAspAspAsnValPhe----- 193
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72989 CAGATGACGACGATGACGACGATGATGACGATGATGATGATGATGATGATGATGATGATGAC 73048
Qy 194 -----AspAspPheThrSerTyrAsnAlaHisLysLysLysGln-----Glu 207
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73049 GAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAA 73108
Qy 208 ArgLysSerArgSerIleAlaAspValArgAsnGluGluGlnAsnIleGlnGlyAsnHis 227
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73109 AATAATAAGAGGATAAAAGGTAACAGAAAGAGGTTCTGAGAAACCATTTGTTGTTAATAAA 73168
Qy 228 ThrGluLeuGlnGluLysSerSerAsnGluAlaThrSerLysGluSerProAlaProLeu 247
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73169 GAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 73228
Qy 248 HisHisArgArgArgMetHisSerArgHisHisLysLeuSerSerArgArgHisGlnGlySer 267
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73229 GATAAAAGAAAAGTATGTAATAAGAAAAGAGG----- 73261
Qy 268 GluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArgHisGlnGlySer 287
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73262 -----AAGAAGAAAGAAAGCGGTG 73282
Qy 288 MetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGly 307
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73283 AATAATTACAACAATAATAAATCAATCAATTAATAATAATAATAATAATAATAATAATAAAG 73342
Qy 308 HisPheGln-----ThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyr 325
```

```
Db 73343 AAGAAAAAGTAATTGGTTAAAGATAAGAAGAAATAGTAGTACCAGTACATGTAATAAATTT 73402
Qy 326 ValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIle----- 342
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73403 ATG-----AATAACTACTATAAATTAATAATAAAAGTTTGTAGTAATATA 73444
Qy 343 -----ValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThr 359
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73445 GCCTCTCATGTATATAAAGGATACAAAT---ATGAGTTCATTAAATACCAGTGAAAT 73501
Qy 360 AsnMetPro---HisLysValHisThrCysHisThr----- 370
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73502 ACTTTACCTTTATATATAATATGGATTGTTCTGCTGTTATAGATATGAAAGAATTATGG 73561
Qy 371 -----SerGlyLeuIleHisLeuGluArgAsnGlu 380
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73562 AAGGATAAGTCGATTTTACCATTAGTTTCCAGGTTTATATATATATATAATAATAA 73621
Qy 381 -----ArgIleHisLeuLysAspIleHisAsnAsp 390
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73622 TTATTGAAGCTAATAACAATTGGTAATAATAATATTCATATGATAGTATATAATAAAC 73681
Qy 391 ArgAsnAlaValLeuArgGluGlyAsnAsnArgSer 402
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73682 AATAATAATAATAATAATAATAATAATAACAATAATAAT 73717

RESULT 14
PFAL29352
LOCUS PFA929352 313050 bp DNA linear INV 29-JAN-2003
DEFINITION Plasmodium falciparum strain 3D7, chromosome 5, segment 2/4.
ACCESSION AL929352 AL844504
VERSION AL929352.1 GI:23504570
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1
AUTHORS Hall N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, P., Davis, P., Dear, P., Dearden, F., Doggett, J., Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Horneby, T., Holroyd, S., Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D., Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N., Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moule, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabinowitz, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and Barrell, B.G.
TITLE Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
JOURNAL Nature 419 (6906), 527-531 (2002)
MEDLINE 2255708
PUBMED 12368867
REFERENCE 2 (bases 1 to 313050)
AUTHORS Devlin, K., Baker, S., Davies, P., Mungal, K., Berriman, M., Pain, A., Hall, N., Bowman, S., Churcher, C., Quail, M. and Barrell, B.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/Projects/P_falciparum.
FEATURES
source
1. 313050
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
```









Residues, Lack of Direct Relation between Consensus Residues and Transmembrane Helices, Expression Patterns of the Transport Protein Genes, and Protein-Protein Interactions with Other Proteins.

5 (bases 1 to 69009)  
Sethuraman,A. and Dolinski,K.J.

Direct Submission  
Submitted (20-DEC-2002) Department of Genetics, Stanford University  
School of Medicine, Saccharomyces Genome Database, Stanford, CA  
94305-5120, USA

Deletion of 16 nt (GATCCTCTCTGATTTT) between 1437732 - 1437747  
Sequence update by submitter

6 (bases 1 to 69009)  
Sethuraman,A. and Cherry,J.M.

Direct Submission  
Submitted (14-OCT-2003) Department of Genetics, Stanford University  
School of Medicine, Saccharomyces Genome Database, Stanford, CA  
94305-5120, USA

Sequence update by submitter  
On Oct 14, 2003 this sequence version replaced gi:27316972.

Sequenced by:  
Stanford DNA Sequence & Technology Center  
855 California Avenue  
Palo Alto, CA 94304, USA

Curated by:  
Saccharomyces Genome Database  
URL: <http://genome-www.stanford.edu/>  
e-mail: [yeast-curator@genome.stanford.edu](mailto:yeast-curator@genome.stanford.edu)

Neighboring Sequence:  
The 5' end of this sequence contains a 1000 bp overlap with GenBank  
Accession Number U33007.

Location/Qualifiers  
1..69009  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/strain="S288c; AB972"  
/db\_xref="taxon:4932"  
/chromosome="IV"  
/map="4R"  
1..12144  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/clones="cosmid 9410"  
2618..38763  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/clones="cosmid 8035"  
38274..69009  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/clones="cosmid 8166"  
68472..69009  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/clones="cosmid 9787"  
521..10327  
/gene="TOM1"  
521..10327  
/gene="TOM1"

/note="YDR457W; Hect-domain-containing nprotein, required for G2/M transition; similar to Rsp5p, ORF YER125w, from SwissProt Accession Number P39940, contains motifs typical of protein kinases; CAI: 0.16"  
/codon\_start=1  
/evidence=not experimental  
/product="Tom1p"  
/protein\_id="AA864910.1"  
/db\_xref="GI:9277738"

gene  
CDS

gene  
CDS

/translation="MVLFTCEKARKEKLAAGKPLVDVLIIDCDTPTFLERIEAIOEM  
DSRRDLYWPIIDLRMDGLLKVAKYIKQDPKKECEVKLVMEAHVDVDCYCKMLK  
FTRLLNLTENFVSSGDMVLLNCPNFTIKLAVMRILAILGQWRFLVAREKIVAHN  
IFGDHNRKTKLALSLSSVMDEGEHFLVDYFKKKYQPKWFKLFTHTYTSND  
FKKSSQCKNNINETSITKVTMTTQELCEHSLQOIFDKGMALLPAEDPWFDSISAV  
AKAFSDGENDIDLRNIIKETKLNATFVNTIIFSPQVSSKLPEDPDYAFNSITDIS  
LSETKIPKELRTDALFTLRCISLKHVWCSDIIRNIGGNSHGLIQILRYIAKTLREA  
THEIDREYNVRFYLLISNADVKPLHESLFAAGLIPTLLEIVSINRCNPYKTLASAT  
LLETIDNSETTEFIENDGFTMLITSVANEIDFTLAHPETWQPKYSVVYYSISFRE  
LAYINSEPTIYVLLVAGLIPYVIDNPKLIGPSAEALLPDWVAICINPGLKQVK  
EKLINLDFLLDADHARILGTGDRSTGYTDIDELARHYDOLKANIVEALCNVIRK  
MPTFRNREFELFTSPKQOKYFRHKNREILTDKEHEPAYWELDKTGMFTFSLV  
FGMSLNGSFSQVPOHLEARDPLAIIIPHENPYEFTSVASINTEVLYLDEKEDY  
AFDMVKLVNDQENLNDLNFSPDRSFLEDRGDSVNSCHSKLRCLAAILNIVTNV  
YIDLTTLSCKRIQIYSPDKRGSFLIKNLKLLFQKCALEEMYIRQHPDSDVITMP  
LPIVDVSGDPLQIYIDDPKKGDKGKITSVKTRNTLQMTLHTPLTQASNTAILRCP  
LRLSHSRNMDLHKLTTEVHFENNVENVIMLKATELGHLPYLVLNNTNFTVFT  
IPKAPNSTEILQITIPAYIFYOKGGVLLVHLIRDLFTMTKI KOLSLDNIYIDES  
NGILTSLCLINALTFYKSMOTETMENVQSIKYVYISDIDDDYINMKALTVPYKVALA  
MILDLKSDSLFKTSRNVPIYVFKQLLSMLKNIITFNVIYTKELYELHMDLIFPIK  
KISLFEQVIGIPGVAANYLITDGDPLPADNSIGLFSBPQWEKYKKLIQEDKSIYFQP  
MQAQQYKGCSEKELDRDTDFNDGLPSPRIFTVLPFKLVNAFAKTLIIQIFTKYDEP  
TEVFAGRILDRILETDLPATLSSLIHLFGILFNEKYIYOKASHMORFIYLBKSL  
KPEHVNTPMFSKALYVYEIILAKSELPHLELSKDVLLARYPLLSNAKVFRIIDPMKOK  
LFDILIRVSDISNFYSALATSRLIIFYSRDELYANNIARSGILSKLVKVGSKQLDK  
INFLSSFLLLTRCFETTENVDALIRAEINKSFARPIGGGDDAVRELTITLSEKAH  
VVMRSPQFIDVLCETARFEPDQDGLVDYSLKRFGLGKDKQNTQASSTESKDIYERT  
GIMHLLLSQMAASEKMDLSEPANSDILPENKQAQLDPSRNPVCAYMIFLLKLVLY  
SSYNCKPEFTTFSRNTYAERPRPTAINPFLYRLDKPVTGDDHKEAKRRVIG  
MLARSVIIGLATVODDRTTKDVKLADPHMNFIRKFAEAIKAIKRNATSSKILLES  
NHLKIDMFRITTSVMYVQAPYLRQLLDSNKVEADQYQCLKVIDLGLPSVITEAMAS  
IDLNPFFSKIFNVAEALNTISRNPFSEHFILDEHDEVEDEVEDKDEIIPMFK  
IDNLGMYDVEDIEEDDDDTSLIGDDDAFVDSNGFEVFEVDSDEDDYDEEDADAR  
SDSENELSSEMQSTADTDVDEYDADGLIINIDQSGDEDEADYDANISSHSS  
SDNHVHREIDSDFLAIPEDIREGILDEAEQRMFGIRIGSADVIADDDVSNDE  
EVENGLDHGNSNRNNDAPPEKKPARIYFAPLIDRAGIASLMKSVFISKPYIREIYH  
ELFYRLCSSKQNRNLDMLNTLFIKLEGIIDQHSLEKVNIISSRAMGHAKTTVRQLP  
SDCTPLTVANOTIEILOSILADSRKLYFLAEHNLIVNKANKSRKEALPDKKLRW  
PLWHLFSLDKRLITDESVMDDLRILQVCTKTLVAVLSTSGNKENSLKXPLPSFD  
EDDLMKILSIIMLDCSTRVFOQTNIILYNSLQGCMSIFTKHLSVLSAISIMSKLS  
ALDGLSREVGTITTGMEINSELLQFTLPSDQAKLKITTVDFLYTHKREBERNV  
KDLQSLYDKMNGGPPWSSLECLSQFEKSOAINTSATILPLIESLMVVCRRSDLSQ  
RNTAVYKEDAKLLDFSKTRVENLPFPPTDAHKLLNOMIRSNPMSGPEALLVKNPK  
VLDPDNKEYFFNAKLKSDNORPKLPITVREOVPLDSYRALFPKTNDEIKNSKLEIT  
FKGSGVDAGVTRREYVQLSRQMPNDYALFLVPDSKTIIPHENRTSGINPEHLSFF  
KPIGMILIGAKRDQCFLCHFSREYVKNILGRVPSLKDMSLDPPYKSLVMILENDI  
TDIIBETFSVETDQGEHKVINLIEGGKDIIVTEANKQDYKVKVEYKQYLSVKEQMD  
NFLVGFYALIISKDLITTFDEQELLLISCLPDIDVDWKNNTVYNTATCKTSYFW  
RAVRSFDAERAKLLOFVTGTSKVLNGLFKELSGVNGVCKFSIHRDPCSSSERLPSHT  
CFNQLNLPVSEYETLRGSLIILAINEGHEGFLA"  
complement (10786..12777)  
/gene="YDR458C"  
complement (10786..12777)  
/gene="YDR458C"  
/note="similar to Saccharomyces cerevisiae hypothetical  
protein on chromosome XIII from GenBank Accession Number  
S49745, CDS number 5; CAI: 0.13"  
/codon\_start=1  
/evidence=not experimental  
/product="Ydr458cp"  
/protein\_id="AAB64934.1"  
/db\_xref="GI:927762"  
/translation="MDHNLDPKTLKVSQRLVLENDVAFANARKPVLKLPKEEKV  
QRLQSSPEASKVTSIQKVVKSAKADRKTKLKKSKLESSSESSTVKDQENVETNK  
RRREIISTDNEAKQIOEKEPKPKKKRKRSSKANKPPSPQSGKATSGADLTSEL  
ETVEELHKKSDDKPRVKELPPELNLKVSNEFLAQLNKLASAAATENYDHSIT

gene  
CDS

DLSSIRIETEPVGPSTGAETRNSESEVENINLEVPQEVKEAEBELTKISTFDNQE  
EDTSRLSKNRPKGRHFIANKTKRGIDIMKPFIAHLFIWLNWGAFLSLICPI  
LFLGWLRQRIQVGVGHEKPLKSLATSAFFQTERVDSVLQAYRPNCLCEPHEGICSS  
FMNVECEPGYEPKSIILTYGIIPPKYCAKDESKVEDELVMKVNVEYLKKNQAQHE  
CGEGENIPESGETETKLYDIRSHRPSHESORENDHWNKVEILKKDDIILWLPLOF  
ETNGREKSKNNNTIYIRSTKRWVLQCHLEGDIQBIYITIKIGSFLITUGVULFLIK  
KIQSTLDNYVQGEQIEKLKVEADIKLQKVNKNGBEPFLFTVQLRATLLSDIPNIKE  
QNNLWAOPTKIMKEQSENIYLLLEENGEIMTCWEWKE"  
/gene="YDR459C"  
/complement (13050. .14174)  
/complement (13050. .14174)  
/complement (13050. .14174)  
/note="similar to Saccharomyces cerevisiae hypothetical  
protein from PIR Accession Number S51289; CAI: 0.12"  
/codon\_start=1

## Alignment Scores:

Pred. No.:	0.841	Length:	69009
Score:	165.50	Matches:	78
Percent Similarity:	35.73%	Conservative:	66
Best Local Similarity:	19.35%	Mismatches:	142
Query Match:	7.65%	Indels:	117
DB:	8	Gaps:	16

US-09-813-329-6 (1-409) x SCD8035 (1-69009)

Qy	51	ValAlaIleLeuAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGlu	70
Db	6086	GTGCTGTTGAAGCTCTAAACACATATTAGTTCGACAGG-----	6124
Qy	71	LeuLysSerLeuLysArgValValAspAsnLeuGlnGlnArgLeuGlyIle-----	87
Db	6125	-----AACAATTTTCTGAGCACCTTCAAAATTTGAAGATCAT	6160
Qy	88	---AsnTyrLeuAspGluPheAspGluPheGlnLysGlu-----TyrGlu	101
Db	6161	GATGAAGTTGAAGACGAGGTTGACGAGTCTGATGAAGAAAGAGATCCCGGATATGTTCAA	6220
Qy	102	AsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAsp-----Glu	117
Db	6221	AACTCAGCGCTG-----GGAATGTATGACGTAGAGGACATTGAA	6259
Qy	118	GluAspAspAspAsp-----GlyAsp-----	124
Db	6260	GAAGACGACGATGACGACACATCATTAATTGGACGCGATGCTATGGCATTTGTTGAT	6319
Qy	125	-----GlyLeuAspSerIleAlaAspAspGluAspAspAspValSerTyrSer	141
Db	6320	AGTGATAACGGATTGGAAGTAGTATTTAGTGACGAAGATGATGATGGGAGAAAGAT	6379
Qy	142	ValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsn	161
Db	6380	GCAGATGATGCTCGTTTCAGATTCGGAGAA---AATGNAATTATCTTCTGAATGCAATCA	6436
Qy	162	AlaHisThrGlyThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAsp	181
Db	6437	AGTACAGCGGACGGACGGATGTGGACTACTGAGGTTGATGATGCAGATGGACTAATTATC	6496
Qy	182	SerAlaSerSerAlaSerAsnAspAsnValPheAspAspPhe-----ThrSer	198
Db	6497	AATATTGACCGCCCTCAGGAGATGATGAAGAAATGGCAGACTATGATGCTTAACATTAGC	6556
Qy	199	TyrAsnAlaHisLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsn	218
Db	6557	CATTCTTCTCAGTGAATAATGAGATGACCGCTCAATGGACGTCATAGAAGTTTACGAC	6616
Qy	219	GluGluGlnAsnIleGlnGlnHisThrGluLeuGlnGlu-----	232
Db	6617	GACGAGCTAAGT---TCTGGATATGATGTTGATTATTAACCGATTATGACGTAGATGAATCT	6673
Qy	233	-----LysSerSerAsnGluAlaThrSerLysGluSer	243
Db	6674	GATTGGGACTCCGGGTTATCCAGCGCTATCTATTCTGATGAAGATAGTGAAGTAGCGAA	6733

Search completed: September 15, 2004, 10:23:34

Job time : 7208 secs

Qy	244	ProAlaProLeuHisHisArgArgMetHisSerArgHisArgHisLeuLeuValArg	263
Db	6734	GATGAGCTATTAACTCCACCAGATGGGAGATTCAGAAGGAGATCGCTGATAGTGAA	6793
Qy	264	LysAlaArg---SerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArg	282
Db	6794	GGCGTGGAGCTAACTGACGACTCGCAA-----	6820
Qy	283	ArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsnGluArg	302
Db	6821	-----GGTGAGTGAAGAAGAT	6838
Qy	303	AsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGly	322
Db	6839	GAT-----AGAGGTGTTTCCGTGGTATTGAACATATATTTTCAAAATGAGAATGAGCCT	6892
Qy	323	LeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIle	342
Db	6893	TTGTTTAGAGTTTCATGACGAAATCGCTCATAGAAACCATCATCGCTCAATCAATCGTACC	6952
Qy	343	ValPheGlnGly-----AspThrProPheLeuGlnCysLeuAsnThrValProThrAsn	360
Db	6953	CATTTTCACTCTGCTATGAGCGGCCCTCTTTAAGCTTATTAAATCTGTGAAGAAGAAAC	7012
Qy	361	MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGlu-----	377
Db	7013	CAAAGCAACCTAATAATCCATAGGCTACAGGATTGGACAACAGTGGAGACGATATT	7072
Qy	378	-----ArgAsnGluArgIleHisLeu	384
Db	7073	TCAGATCAGGTGACTGTTGCTGTTGCTGAGGCTCTCGCCCAAGATCACATCATCTTCACTTT	7132
Qy	385	LysAspIle	387
Db	7133	TCGGAAGTT	7141



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 02:17:51 ; Search time 634 seconds  
(without alignments)  
2740.557 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTAETLKPFITPTSANDGF.....DRNAVLRGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09813329/runat\_13092004.102752.21656/app\_query.fasta\_1.583  
-DB=N-Geneseq 29Jan04 -QWMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09813329@cgn 1.1.885 @runat\_13092004.102752.21656 -NCPU=6 -ICPU=3  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2003as: \*  
8: geneseqn2003bs: \*  
9: geneseqn2003cs: \*  
10: geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
-----						
1	2162	100.0	2148	6	ABK11680	Abk11680 DNA encod
2	2082.5	96.3	1221	6	ABK11679	Abk11679 DNA encod
3	2045	94.6	2166	6	ABK11678	Abk11678 DNA encod
4	1656	76.6	978	4	ABL21473	ABL21473 Drosophil
5	1483.5	68.6	3324	4	ABL21472	ABL21472 Drosophil
6	151	7.0	2616	6	ABes4139	ABes4139 Mouse cdn
7	145.5	6.7	11528	4	ABL15995	ABL15995 Drosophil
8	141	6.5	1287	6	ABQ76511	ABQ76511 C. albica

c	9	140	6.5	567	3	AAA29550	Aaa29550 HIV codon
	10	138.5	6.4	3579	3	AAA70099	Aaa70099 Plasmodiu
	11	138.5	6.4	6240	4	ABL06443	AbL06443 Drosophil
	12	138	6.4	1176	6	ABL51020	AbL51020 Human EDA
	13	138	6.4	1176	8	ACD07906	AcD07906 DNA encod
	14	138	6.4	1574	6	ABL51009	AbL51009 Human EDA
	15	138	6.4	1574	8	ACD07895	AcD07895 cDNA enco
	16	138	6.4	5307	7	ACC57577	Acc57577 Polynucle
	17	138	6.4	5307	7	ACC57910	Acc57910 Human EDA
	18	138	6.4	5307	9	ADC35223	AdC35223 Human EDA
	19	137	6.3	1357	3	AZ51257	Az51257 Human RNA
	20	137	6.3	2635	4	ABL20713	AbL20713 Drosophil
c	21	137	6.3	4183	4	AAK74968	AaK74968 Human imm
c	22	137	6.3	4183	4	AAK75364	AaK75364 Human imm
	23	137	6.3	5276	4	ABL20712	AbL20712 Drosophil
	24	137	6.3	6796	4	ABL17763	AbL17763 Drosophil
	25	136	6.3	40059	4	ABL15994	AbL15994 Drosophil
	26	135	6.2	1626	4	AAS35027	AaS35027 cDNA enco
	27	135	6.2	1626	4	ABK43159	AbK43159 cDNA enco
	28	135	6.2	1626	9	ADC46185	AdC46185 Human neo
	29	135	6.2	2075	4	ABL21925	AbL21925 Drosophil
c	30	135	6.2	4248	4	ABL21924	AbL21924 Drosophil
	31	134	6.2	1401	7	ACC60802	Acc60802 Gene sequ
	32	133	6.2	2614	4	AAH16318	AaH16318 Human cdn
	33	133	6.2	3489	3	AAA30290	Aaa30290 Kaposi's
	34	133	6.2	3489	4	AAF82901	AaF82901 Nucleotid
	35	133	6.2	3489	6	ABA93487	Aba93487 Kaposi's
	36	133	6.2	7503	3	AA70206	Aa70206 Plasmodiu
c	37	133	6.2	32207	2	AAV73805	Aav73805 KSHV LOR
c	38	133	6.2	137507	2	AAV19941	Aav19941 KSHV long
	39	130.5	6.0	3615	4	ABL03529	AbL03529 Drosophil
	40	129	6.0	4599	2	AAT27052	Aat27052 RRP3 telo
	41	128.5	5.9	1516	4	ABL17433	AbL17433 Drosophil
	42	128.5	5.9	3100	7	AA51702	Aa51702 Human nuc
	43	128.5	5.9	3578	4	ABL17432	AbL17432 Drosophil
	44	128.5	5.9	4522	2	AAV74350	Aav74350 Staphyloc
	45	128	5.9	2619	6	ABZ32134	AbZ32134 Candida a

ALIGNMENTS

RESULT 1	ABK11680	ABK11680 standard; DNA; 2148 BP.
ID	ABK11680	
XX	XX	ABK11680;
AC	ABK11680;	
XX	XX	05-JUN-2002 (first entry)
DT	05-JUN-2002	
XX	XX	DNA encoding tumour necrosis factor variant 2 (TNFv2).
DE	DE	Tumour necrosis factor; TNF; pesticide; crop-damaging insect;
KW	KW	farm animal insect; epithelial morphogenesis; cell-matrix adhesion;
KW	KW	x-linked anhidrotic (hypohidrotic) ectodermal dysplasia;
KW	KW	x-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;
KW	KW	sparse hair; sweat gland aberration; endotoxin shock; inflammation;
KW	KW	haemorrhagic necrosis of tumour; cytotoxicity; TNFv2;
KW	KW	obesity-linked insulin resistance; gene; ds.
OS	OS	Drosophila melanogaster.
FH	FH	Key Location/Qualifiers
FT	FT	CDS 634..1863
FT	FT	/*tag= b
FT	FT	/product= "TNFv1"
FT	FT	/note= "Tumour necrosis factor variant 1"
FT	FT	misc_difference 634..1860
FT	FT	/*tag= d
FT	FT	/note= "Specifically claimed in claim 22"
FT	FT	634..789
FT	FT	/*tag= a
FT	FT	790..1860
FT	FT	/*tag= c

FT XX /label= mature\_TNFv1  
 PN XX US2002012968-A1.  
 PD XX 31-JAN-2002.  
 PF XX 20-MAR-2001; 2001US-00813329.  
 XX XX 21-MAR-2000; 2000US-0190816P.  
 PR XX (CARR/) CARROLL P M.  
 XX (CHEN/) CHEN J.  
 PA (RAMA/) RAMANATHAN C S.  
 PA (XIAO/) XIAO H.  
 PA (GUAN/) GUAN B.  
 PA (BOWE/) BOWEN M A.  
 XX  
 PI Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;  
 XX  
 DR WPI; 2002-195121/25.  
 DR P-PSDB; AAU77718.  
 XX  
 PT New Drosophila tumor necrosis factor molecule, useful in controlling  
 PT agriculturally important pests, e.g. comprises modifying the growth,  
 PT feeding or reproduction of crop-damaging insects or insects of farm  
 PT animals.  
 XX  
 PS Claim 2; Fig 3A-C; 11pp; English.  
 XX  
 CC The invention describes an isolated tumour necrosis factor polypeptide  
 CC (TNF). The polypeptide and polynucleotide are useful in controlling  
 CC agriculturally important pests, particularly by modifying the growth,  
 CC feeding and/or reproduction of crop-damaging insects or insects of farm  
 CC animals. The polypeptide and polynucleotide are useful for modulating  
 CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus  
 CC the polypeptide and polynucleotide may be useful for treating,  
 CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal  
 CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-  
 CC like disorders, e.g. sparse hair, abnormal or missing teeth or sweat  
 CC gland aberrations in animals (e.g. insects and potentially humans),  
 CC endotoxic shock, inflammation, haemorrhagic necrosis of tumours,  
 CC cytotoxicity and obesity-linked insulin resistance, all of which involve  
 CC TNF molecules. This sequence encodes the drosophila melanogaster tumour  
 CC necrosis factor variant 2 (TNFv2) protein, described in the invention  
 XX  
 SQ Sequence 2148 BP; 645 A; 510 C; 506 G; 487 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3.52e-179 Length: 2148  
 Score: 2162.00 Matches: 409  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-813-329-6 (1-409) x ABK11680 (1-2148)  
 Qy 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe 20  
 Db 634 ATGACTGCCGAGACCTTCAGCCGTTTATACGCCAACGAGTGCCACGATGATGTTT 693  
 Qy 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40  
 Db 694 CCGGCCAAAGCGACAGCAGCGGCCGCCAGCGACGCCAGCCCGCAGCTGATCCCGCTG 753  
 Qy 41 ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaIleThrIleTrpGln 60  
 Db 754 GTTTTGGGGTTCATCGGTCTGGGGTCTGGTCTGGCCATCTCGCACTAACGATCTGGCAG 813  
 Qy 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn 80  
 Db 814 ACAACGGTGTATCGCATCTGGACAGGAGCTGAAGAGCCTGAAGCGAGTGTGCGTAAAT 873

RESULT 2  
 ABK11679  
 ID ABK11679 standard; DNA; 1221 BP.  
 XX  
 AC ABK11679;

Qy 81 LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100  
 Db 874 CTCAGCAGCGTTTGGGCATAAACTATCTGGACGAGTTCGACGAGTTCCAAAGAGGAGTAC 933  
 Qy 101 GluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluAspAsp 120  
 Db 934 GAGAAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCTTCACGGATGAGGAGCAGC 993  
 Qy 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer 140  
 Db 994 GACGATGGCGATGGTCTGGATTCATTTGGGACGACGAGACGACGCTTACGTATATAGC 1053  
 Qy 141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn 160  
 Db 1054 TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAATAAATCAAC 1113  
 Qy 161 AsnAlaHisThrGlyThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180  
 Db 1114 AATGCACATACCGGCACCCACCGCCCATCTGAGACCATCTGAGGGCGGAGGGCGAGACG 1173  
 Qy 181 AspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyrAsn 200  
 Db 1174 GACAGTGCATCTCTCAGCCTCAATGATGACATGTTGTCGATGACTTTACCAAGCTACAA 1233  
 Qy 201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGlu 220  
 Db 1234 GCCCAAAAAAGAGCAGGAGAGAAAAATCTCGCTCGATTGCGGATGTACGCAATGAGGAG 1293  
 Qy 221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAlaThrSer 240  
 Db 1294 CAGATATTTCAGGAAATCACAAGAGCTTCAGGAAAGTCAATCCAAATGAGGCAACTTCC 1353  
 Qy 241 LysGluSerProAlaProLeuHisHisArgArgMetHisSerArgHisArgHisLeu 260  
 Db 1354 AAAGAGAGCCTGCAACCACTTCCACCGTCGACAGATGTCATTCGCGCATCGCCACCTC 1413  
 Qy 261 LeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSer 280  
 Db 1414 CTAGTCGCGAAAGCCAGATCCGAGGACTCGAGGCCAGCAGGCCCAATTTCCACTTGAGCAGC 1473  
 Qy 281 ArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn 300  
 Db 1474 AGCGCGCGCTCAACAGGAAGTATGGCTACCATGGAGATATGTACATAGGAAATGATTAAC 1533  
 Qy 301 GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn 320  
 Db 1534 GAGAGAAACTCTTATCAGGGACACTTTCAAACGCGCGATGGCGTCTTGACGCTGACCAAT 1593  
 Qy 321 ThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAsnGly 340  
 Db 1594 ACAGGCTTATATTACGTATACGCCAGATATGCTACAACTCGACGACGACGACGACG 1653  
 Qy 341 PheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsn 360  
 Db 1654 TTTATCGTCTTCAAGGAGACACTCCATCTCTGCGAGTGTCTTGAACACGCTGCCCAAC 1713  
 Qy 361 MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGlu 380  
 Db 1714 ATGCCACATAAAGGTGCACACCTGCCACACGATGCGATGCTGATCCACCTGGAAACGAA 1773  
 Qy 381 ArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsn 400  
 Db 1774 AGGATCCATCTGAAGGACATTCACACGATCGCAATGCAGTTCCTGCGGGAGGAAACAC 1833  
 Qy 401 ArgSerTyrPheGlyIlePheLysVal 409  
 Db 1834 CGAAGACTACTTGGCATCTTCAAGGTG 1860

XX 05-JUN-2002 (first entry)  
XX DNA encoding tumour necrosis factor variant 1 (TNFv1).  
DE Tumour necrosis factor; TNF; pesticide; crop-damaging insect;  
XX farm animal insect; epithelial morphogenesis; cell-matrix adhesion;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;  
KW sparse hair; sweat gland aberration; endotoxigenic shock; inflammation;  
KW haemorrhagic necrosis of tumour; cytotoxicity; TNFv1;  
KW obesity-linked insulin resistance; gene; ds.  
XX Drosophila melanogaster.  
OS  
XX  
XX Key Location/Qualifiers  
XX CDS 1..1221  
XX /\*tag= b  
XX /product= "TNFv1"  
XX /note= "Tumour necrosis factor variant 1"  
XX misc\_difference 1..1218  
XX /\*tag= d  
XX /note= "Specifically claimed in claim 18"  
XX sig\_peptide 1..156  
XX /\*tag= a  
XX mat\_peptide 157..1218  
XX /\*tag= c  
XX /label= mature\_TNFv1  
XX  
XX US2002012968-A1.  
XX  
XX 31-JAN-2002.  
XX  
XX 20-MAR-2001; 2001US-00813329.  
XX  
XX 21-MAR-2000; 2000US-0190816P.  
XX (CARR/) CARROLL P M.  
XX (CHEN/) CHEN J.  
XX (RAMA/) RAMANATHAN C S.  
XX (XIAO/) XIAO H.  
XX (GUAN/) GUAN B.  
XX (BOWE/) BOWEN M A.  
XX  
XX Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;  
XX  
XX WPI, 2002-195121/25.  
XX  
XX New Drosophila tumor necrosis factor molecule, useful in controlling  
XX agriculturally important pests, e.g. comprises modifying the growth,  
XX feeding or reproduction of crop-damaging insects or insects of farm  
XX animals.  
XX  
XX Claim 2; Fig 2A-B; 119pp; English.  
XX  
XX The invention describes an isolated tumour necrosis factor polypeptide  
XX (TNF). The polypeptide and polynucleotide are useful in controlling  
XX agriculturally important pests, particularly by modifying the growth,  
XX feeding and/or reproduction of crop-damaging insects or insects of farm  
XX animals. The polypeptide and polynucleotide are useful for modulating  
XX epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus  
XX the polypeptide and polynucleotide may be useful for treating,  
XX ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal  
XX dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-  
XX like disorders, e.g. sparse hair, abnormal or missing teeth or sweat  
XX gland aberrations in animals (e.g. insects and potentially humans),  
XX endotoxigenic shock, inflammation, haemorrhagic necrosis of tumours,  
XX cytotoxicity and obesity-linked insulin resistance, all of which involve  
XX TNF molecules. This sequence encodes the drosophila melanogaster tumour  
XX necrosis factor variant 1(TNFv1) protein, described in the invention  
XX  
XX Sequence 1221 BP; 349 A; 323 C; 312 G; 237 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.54e-172 Length: 1221  
Score: 2082.50 Matches: 400  
Percent Similarity: 96.39% Conservatives: 0  
Best Local Similarity: 96.39% Mismatches: 0  
Query Match: 96.32% Indels: 15  
DB: 6 Gaps: 2  
US-09-813-329-6 (1-409) x ABK11679 (1-1221)  
Qy 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAanAspAspGlyPhe 20  
Db 1 ATGACTCCGAGACCCCTCAAGCCGTTTATACGCCAACGAGTGCACACGATGATGGTTTT 60  
Qy 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuLeuProLeu 40  
Db 61 CCGGCCAAGGACGACGACGCGCCGAGGACGACCCGCGCAGCTGATCCCCCTG 120  
Qy 41 ValLeuGlyPheLeuGlyLeuValValAlaLeuAlaLeuLeuThrIleTrpGln 60  
Db 121 GTTTTGGGGTTTCATCGGCTCTGGGCTGGTGGTTCGTCATTCGCACTTCGCACTAAGCATCTGGCAG 180  
Qy 61 ThrThrArgValSerHisLeuAspLysLeuLeuLysSerLeuLysArgValValAspAsn 80  
Db 181 ACAACGCGTGTATCGCATCTGGACAAGGAGCTGAAGAGCCTGAAGCGAGTCTGCGATAAT 240  
Qy 81 LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100  
Db 241 CTCGAGCAGCGTTTGGGCATTAACCTATCTGGACGAGTTCGACGAGTTCCTCAAAAGGAGTAC 300  
Qy 101 GluAsnAlaLeuLeuAspTyrProLysLysValAspGlyLeuThrAspGluGluAspAsp 120  
Db 301 GAGAAATGCCCTCATCGACTATCAAAAAAGGTGGTGGCTCTCACGGATGAGGAGGACGAC 360  
Qy 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspAspValSerTyrSer 140  
Db 361 GACGATGCGATGGTCTGGATTCCATTTGCGGACGACGAGGACGAGCGAGCTAGCTATAGC 420  
Qy 141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn 160  
Db 421 TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCCGATATGTTAAATAAATCAAC 480  
Qy 161 AsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180  
Db 481 AATGCACATACCGCACCGCCACATCTGAGACCACTCTGAGGCGGAGGCGGCGAGACG 540  
Qy 181 AspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheThrSerTyrAsn 200  
Db 541 GACAGTGCATCCTCAGCCTCAAAATGATGACAATGTGTTCGATGACTTTTACGAGTACAAT 600  
Qy 201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu 220  
Db 601 GCCCACAAAAGAGGAGGAGGAGAAAATCTCGCTCGATTGCCGATGTACCAATGAGGAG 660  
Qy 221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlnLysSerSerAsnGluAlaThrSer 240  
Db 661 CAGAAATATTCAAGAAATCACACAGAGCTTCAGGAAAAGTATCCCAATGAGGCAACTTCC 720  
Qy 241 LysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHisLeu 260  
Db 721 AAAAG-----AGAAATGCAATTTCCCGCCATCGCCACCTC 753  
Qy 261 LeuValArgLys-----AlaArgSerGluAspSerArgProAlaAla 274  
Db 754 CTAGTCCGCAAGGTGAATCTTCTTTCAGCCAGATCCGAGGACTCCGAGCCAGCAGCC 813  
Qy 275 HisPheHisLeuSerSerArgArgArgHisGlnGlnLysSerMetGlyTyrHisGlyAspMet 294  
Db 814 CATTTCCACTTGAGCAGCGCGCTGACCAAGGAAGTATGGCTACCATCGAGATATG 873  
Qy 295 TyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGly 314  
Db 874 TACATAGGAATGATACGAGAGAAACTCTTATCAGGGACACTTTTCAACGCGCGATGGC 933



QY 181 AppSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheThrSerTyrAsn 200  
Db |||||  
1192 GACAGTGCATCTCAGCCTCAATGATGACAATGTGTTTCGATGACTTTACCGACTCAGAT 1251  
QY 201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu 220  
Db |||||  
1252 GCCCTCAAAAGAACGAGAGAGAAAATCTCGCTCGATTTGCCGATGTACGCAATGAGGAG 1311  
QY 221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAlaThrSer 240  
Db |||||  
1312 CAGATATTCAAGGAATACACAGAGCTTCAGGAAAAGTCAATCCATGAGGCGACTTCC 1371  
QY 241 LysGluSerProAlaProLeuHisHisArgArgMetHisSerArgHisLeu 260  
Db |||||  
1372 AAAGAGAGCCCTGCAGCATTTCACCTCGTCGAGCAATGCATTCGCCCATCGCCACCTC 1431  
QY 261 LeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSer 280  
Db |||||  
1432 GTAGTCCGCAAGCCAGATCCGAGACTCGAGGCCAGCAGGCCCATTTCCACTTGAGCAGC 1491  
QY 281 ArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn 300  
Db |||||  
1492 AGCGCGCTCACCAGAAAGTATGGCTACCATCGAGATATGTACATAGAAAATGATAGG 1551  
QY 301 GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn 320  
Db |||||  
1552 GAGAGATGCTCTTATCAGGACACTTTCAAACGCGCGATGGCGTATTGACGGTGACCAAT 1611  
QY 321 ThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAsnGly 340  
Db |||||  
1612 CGAGGCCCTATTACGTATACGCCAGATATGGGGCTACAACTCGCAGATCCAGAACGGA 1671  
QY 341 PheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsn 360  
Db |||||  
1672 TTTATCGTCTTCAAGGAGACACTCCATTCCTGCGAGTCTTGAACACCGGTGCCCAAC 1731  
QY 361 MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGlu 380  
Db |||||  
1732 ATGCCACATAAGGTGCACACCTGCCACACGAGTGGTCTGATCCACTCGAACGAAACGAG 1791  
QY 381 ArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsn 400  
Db |||||  
1792 AGGATCCATCTGAAGGACATTCACAACGATCGCAATGCGATCTCTCGGGGAGGAAACAAC 1851  
QY 401 ArgSerTyrPheGlyIlePheLysVal 409  
Db |||||  
1852 CGAAGCTACTTGGCATCTTCAAGGTG 1878

## RESULT 4

ABL21473  
ID ABL21473 standard; DNA; 978 BP.  
XX  
AC ABL21473;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 15892.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical; gene; db.  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
FA (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.  
XX  
XX Claim 1; SEQ ID NO 15892; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
XX sequences (ABU01840-ABU16175) and the encoded proteins (ABBS7737-  
XX ABBS72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 978 BP; 289 A; 261 C; 238 G; 190 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.19e-135 Length: 978  
Score: 1656.00 Matches: 310  
Percent Similarity: 98.11% Conservative: 1  
Best Local Similarity: 97.79% Mismatches: 0  
Query Match: 76.60% Indels: 6  
DB: 4 Gaps: 1

US-09-813-329-6 (1-409) x ABL21473 (1-978)

QY 99 GluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluGlu 118  
Db |||||  
25 CAGTACGAGAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCTTCAGGATGAGGAG 84  
QY 119 AspAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSer 138  
Db |||||  
85 GACGACGACGATGCGATGCTGGATTCATTGCGGACGACGAGGACGAGCTTAGC 144  
QY 139 TyrSerSerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLys 158  
Db |||||  
145 TATAGTCTCTGGATGATGTTGGCGAGACTACGAGGACTACACCGATATGTTAAATAAA 204  
QY 159 LeuAsnAsnAlaHisThrGlyThrProThrSerGluThrAlaGluGlyGluGly 178  
Db |||||  
205 CTCACAATGCATATACCGGACCCACCGCCACATCTGAGACCACTGCTGAGGGCGAGGC 264  
QY 179 GluThrAspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSer 198  
Db |||||  
265 GAGCGGACAGTGCATCTCAGCCTCAATGATGACANTGTGTTTCGATGACTTTTACCAGC 324  
QY 199 TyrAsnAlaHisLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsn 218  
Db |||||  
325 TACAATGCCCAAAAAGAACGAGGAGAGAAAATCTCGCTCGATTCGCGATGTACGCAAT 384  
QY 219 GluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAla 238  
Db |||||  
385 GAGGAGCAGATAATTCAAGGAAATCACACAGAGCTTCAGGAAAAGTCAATCCAATGAGGCA 444  
QY 239 ThrSerLysGluSerProAlaProLeuHisHisArgArgMetHisSerArgHisArg 258  
Db |||||  
445 ACTTCAAAGAGAGCCCTCGCACCTTCCACCGCTCGCAGAAATGCAATTCCTCCGCAATCGC 504  
QY 259 HisLeuLeuValArgLys-----AlaArgSerGluAspSerArgPro 272  
Db |||||  
505 CACCTCTAGTCCGCAAAAGGTGAATCTCTTCTTTCAGCCAGATCCGAGGACTCGAGGCCA 564  
QY 273 AlaAlaHisPheHisLeuSerSerArgArgHisGlnGlySerMetGlyTyrHisGly 292  
Db |||||  
565 GCAGCCCATTTCCACTTGAGCAGCAGCGCGCTCACCAAGGAAGTATGGGCTACCATGGA 624





Db 1402 CAGGCAACCTAATGAAGTCACCTTTCTGAATTTCTCA -GGTCACCATGAGATATGTACAT 1344  
 Qy 296 eGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValle 316  
 Db 1343 AGGAAATGATAACGAGAGAAACTTTATCAGGGACACTTTCAACGCCGCGATGGCGTCTT 1284  
 Qy 316 uThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHi 336  
 Db 1283 GACGGTGACCAATACAGGCCTATATTAGCTATACGCCAGATATGCTACAACTCGCA 1224  
 Qy 336 sAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnTh 356  
 Db 1223 GCACAGAACCGAATTTATCGCTTTCAAGGAGACACTCCATCTCTGCGAGTCTTGAACAC 1164  
 Qy 356 rValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLe 376  
 Db 1163 GGTGCCCAACCAATGCCACATNAGTGCACACCTGCCACAGAGTGCTGTGATCCACCT 1104  
 Qy 376 uGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuAr 396  
 Db 1103 GGAACGAACGAGAGGATCCATCTGAAGGACATTCACAACTGCAATGCAATGCGTCTGCG 1044  
 Qy 396 gGluGlyAsnAsnArgSerTyrPheGlyIlePheIleVal 409  
 Db 1043 GGAGGAAACCAACCGAAGTACTTTGGCATCTTCAAGGTG 1004  
 RESULT 6  
 ID ABS54139  
 XX ABS54139 standard; cDNA; 2616 BP.  
 AC ABS54139;  
 XX  
 DT 25-NOV-2002 (first entry)  
 XX  
 DE Mouse cDNA encoding Zis-SR protein.  
 XX  
 KW Mouse; ss; gene; Zis-SR; neuroendocrine phenotype; diabetes;  
 KW Parkinson's disease; Alzheimer's disease; neurodegenerative disease;  
 KW zinc finger splicing with extended Ser-Arg domain; secretory pathway;  
 KW zinc finger protein.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT 23..1015  
 CDS /\*tag= a  
 FT /product= "Zis-SR"  
 FT  
 FT  
 FN WO200261082-A2.  
 XX  
 XX 08-AUG-2002.  
 XX  
 XX 29-JAN-2002; 2002WO-CA000101.  
 XX  
 XX 29-JAN-2001; 2001US-0264296P.  
 XX  
 XX (UYSH ) UNIV SHERBROOKE.  
 PA  
 XX  
 XX Day R;  
 PI  
 XX WPI; 2002-682683/73.  
 DR  
 DR P-ESDB; ABG32871.  
 XX  
 XX New Zis-SR nucleic acid molecules and polypeptides, useful for restoring  
 PT or increasing the secretory properties of a cell, or for treating  
 PT diseases or conditions associated with a loss of function, e.g. diabetes  
 PT or Parkinson's disease.  
 XX  
 XX Claim 1; Fig 4; 70pp; English.  
 PS  
 XX The invention relates to an isolated nucleic acid molecule, Zis-SR,  
 CC encoding a protein involved in the secretory pathway in a cell ( or its  
 CC homologue or variant) or nucleic acid molecules that hybridise under high

CC stringency condition to the Zis-SR nucleic acid. Also included are an  
 CC isolated polypeptide involved in the formation of secretory granules in  
 CC cells comprising the amino acid sequence spanning amino acids 243-310 of  
 CC the Zis-SR protein, restoring the neuroendocrine differentiation of a  
 CC cell using the nucleic acid molecule or polypeptide cited above,  
 CC identifying a gene and/or protein involved in inducing regulated  
 CC secretion comprising a comparison at the molecular level of a secretion-  
 CC defective cell line under conditions that restore differentiation of the  
 CC secretion-defective cell, such that secretion is restored, and the  
 CC secretion-defective cell line in the absence of the conditions cited.  
 CC Also included are modulating the secretory properties of a cell  
 CC comprising modulating the activity and/or level of Zis-SR and an assay to  
 CC identify a modulator of regulated secretion in a cell comprising an  
 CC assessment of a biological activity of Zis-SR, its part or derivative in  
 CC the presence of a candidate agent, where a modulator of regulated  
 CC secretion is selected when the biological activity of Zis-SR, its part or  
 CC derivative is measurably different in the presence of the candidate  
 CC compound as compared in its absence. The nucleic acid molecules or  
 CC polypeptides are useful for restoring or increasing the secretory  
 CC properties of a cell, for regulating neuroendocrine phenotype, and for  
 CC long term therapies to treat diseases or conditions associated with a  
 CC loss of function, e.g. diabetes, neurodegenerative diseases such as  
 CC Alzheimer's disease or Parkinson's disease. The assay is useful for  
 CC screening compounds for treating diseases or conditions associated with a  
 CC defect in the regulated secretory pathways in cells. The nucleic acid  
 CC molecules can also be used to locate gene regions associated with genetic  
 CC diseases. The present sequence encodes mouse Zis-SR (zinc finger splicing  
 CC with extended Ser-Arg domain)  
 XX  
 SQ Sequence 2616 BP; 834 A; 458 C; 542 G; 782 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 0.00257 Length: 2616  
 Score: 151.00 Matches: 93  
 Percent Similarity: 33.84% Conservative: 63  
 Best Local Similarity: 20.17% Mismatches: 156  
 Query Match: 6.98% Indels: 150  
 DB: 6 Gaps: 17

US-09-813-329-6 (1-409) x ABS54139 (1-2616)

Qy 20 PheProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuIlePro 39  
 Db 95 TTTGCTAGGAGAACCCAGCTGTAACAGATGTGTGTCGAGAAAAGACACTGAGGCCAAGATG 154  
 Qy 40 LeuValLeuGlyPheIleGlyLeuGlyLeuValValAla-----IleLeu 54  
 Db 155 ATGAAAGCTGGGGGAACAGAAATAGGAAAGACACTGGCAGAGAACGCCGGGCTTATTT 214  
 Qy 55 AlalaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLeu 74  
 Db 215 AGTGCCCAATGATTGGCAATGCCAAACTTGCAGTAATGTGAATTGGGCTAGAGATCAGAG 274  
 Qy 75 LysArgValValAsp-----AsnLeuGlnGlnArgLeuGly----- 86  
 Db 275 TGTAAACATGTGTAATACTCCAAAGTATGCTAAATATAGAAAGAAAGACAGGATGGAGGT 334  
 Qy 87 -----IleAsnTyrLeuAsp-----GluPhe 93  
 Db 335 GGTTTAAATGAAAGAGAGAATGTTGAATACATAGAAAGAGAGAAGAACTCGATGGGAATAT 394  
 Qy 94 AspGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGly 113  
 Db 395 GATGAGTTTGAGCTGATAAAAGAAAAA-----TACAGGGGGAAGCAGTGGC 442  
 Qy 114 -----LeuThrAspGluGluAspAspAspGlyAspGlyLeuAspSerIleAla 130  
 Db 443 CCTGTCATCTATATTAAAGGAAGTTGAAGATAAAGAGTCAGAGGAGAGA----- 493  
 Qy 131 AspAspGluAspAspAspValSerTyrSerValAspAspValGlyAlaAspTyrGlu 150  
 Db 494 GAGGATGAGGATGAAGATCTTCTTAAGTATATAAGCTAGAT----- 532

151	AspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThrGlyThrThrProThrSer	170
532	-----	532
171	GluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAsp	190
533	-----GAGACGAGGATGAAGATGATGCTGATCTCTCAAAATATATAAT	574
191	AsnValPheAspAspPheThrSerTyrAsnAlaHisLysLysGlnGluArgLysSer	210
575	-----CTTGATGCCGCGAAGAGATAGTAAACAAAAGAAAGCAATAGGCGGAGC	628
211	ArgSerIleAlaAspValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeu	230
629	CGCTCA-----	634
231	GlnGluLysSerSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArg	250
635	AAGTCACGATGCTCTCACTCAAGGCTTCATCACGCTCATCTCCCCCTCAAGTTCAAGG	694
251	ArgArgMetHisSerArgHisArgHisLeuLeuValArgLysAlaArgSer	267
695	TCCAGGTCCAGTCCCGTTCAGAGCTTCCAGTTCGCAGTCCAGGTCCTCACTCCCGT	754
268	-----GluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArgHisGln	285
755	TCCAGAGAACATTCAGATCCCGTGGTTCGAAATCAAGATCCAGTCCAGGTCACACAGG	814
286	GlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsn-----Asp	299
815	GGCTCTTCTCCCMAGAAAAGATCTTATTTCGAGTTCTTCGTCACTCTCGTAAAGAC	874
300	AsnGluArgAsnSerTyrGlnGlyHisPhe-----GlnThrArgAspGlyVal	315
875	AGGAAGAGAGTCGCTAGACC-TTCTTTCACCAGCTGTTCCGCAAAAACACGACGAG	933
316	LeuThrValThrAsnThrGlyLeuTyrTyrValTyr-----	327
934	ATCAGGTCACCCGAAAGGCACACAGGTCGCTCCGGATCAACACATTCTCGTTCCCG	993
328	-----AlaGlnIleCysTyrAsnAsnSerHis-----	336
994	TTCAAGTTCAAAAAGAAATATGATTTAAATTTACATCTTTAAAAACATTGAGTACA	1053
336	-----	336
1054	GTGCATGAAGCATATTTTTAGGAAGTTGATGCTCTATTGTCAGAAGTACTACATCTG	1113
337	-----AspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeu	354
1114	CTAGTAGAGGTGCATGCTTTATTGCTTTTCAAAACAACAACACTGTGTTTATTGTT	1170
355	AsnThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIle	374
1171	---GAAACTGAAGTAAATTGCATTTTAAAGCCATATACTGTTCGCCAGATAAATGCTCTG	1227
375	HisLeuGluArgAsnGluArgIleHisIleLysAspIleHisAsnAspArgAsnAlaVal	394
1228	TTCAATTATTACACCACTTTCCTTCATTTAAAA-----CATCTCAGCAATACAGAGTA	1281
395	Leu 395	
1282	CTT 1284	

Qy	I26	LeuaspSerIleAlaaspGluaspGluaspAspValSerySerSerValaspAspVal	145
Db	1574	-----GATGATGATGACGAGC-----GACGCACATGGAGCGTGCTG	1609
Qy	146	GlyAlaaspTyrGluaspTyrThrAspMet-----LeuAsnLysLeuAsnAsnAlahis	163
Db	1610	:::     :::     :::     :::     :::     :::	1669
Qy	164	ThrGlyThrThrProthrSerGluuthrThrAlaGluGlyGluGlyGluuthrAspSerAla	183

	Qy	126	LeuAspSerIleAlaAspGluAspAspValSery-SerSerValAspAspVal	145
	Db	1574	-----GATGATGATGACGAGC-----GACGACATGGAGCTGTG	1609
	Qy	146	GlyAlaAspTyrGluAspTyrThrAspMet-----LeuAsnLysLeuAsnAsnAlahis	163
	Db	1610	:::     :::     :::     :::     :::     :::	1669
	Qy	164	ThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAla	183

Db 1670 TCAGCGGCGTGAATGATCCCAACAAACAGCAGCATCATCAGGC-----GGA 1717

QY 184 SerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyrAsnAlaHisLys 203

Db 1718 GCAGCAACATCCGACGACGATAAACTGATGTGATGCGATGTTGGGGGGCTGCCT 1777

QY 204 LysLysGlnGluArg-----LysSerArg 211

Db 1778 AAACCGCAGCATCATCTTTGTCAGAAATCTGCTCGCGCGCAGCGATGACAAATCCCGAG 1837

QY 212 SerIleAlaAspValArgAsnGluGlnAsnIleGlnGlyAsnHisThrGluLeuGln 231

Db 1838 CATTTGAGTGATCATGACAAATGAATCTTTCAACAGCATCGATTTTCGATGCCGGAATACACA 1897

QY 232 GluLysSerSerAsnGluAlaThrSerLysGluSerProAlaPro----- 246

Db 1898 ATTGCGCGCAGCAAGAGCAGCAGCAGCAACATCCACCGCGGATGATTCGGTGGATCC 1957

QY 247 -----LeuHisHisArg-ArgMe 253

Db 1958 GCGCATCAACAGCAATTTGGCAGCATTTCTTCAACAAATCTGCTGTCGATTTCCAATGCAGCT 2017

QY 253 tHisSerArgHisArgHisLeuLeuValArgLysAlaArgSerGluAspSerArgProAl 273

Db 2018 TCGGAGTCGGTCAGTAAGCTCTTTTCGTCATCAACAGCG---GCTCTAAGTCCACGCCCTCC 2074

QY 273 aAla-HisPheHisLeuSerSerArgArgArg-----HisGlnGlySerMetGlyT 290

Db 2075 AAGTCCGGTTTCCACTCCGGCTCCCGCCGACAAATCGGATGCGATCTCGGCGCTCGCTT 2134

QY 290 Yr 290

Db 2135 AC 2136

RESULT 8

ABQ76511

ID ABQ76511 standard; cDNA; 1287 BP.

XX AC ABQ76511;

XX 21-NOV-2002 (first entry)

XX DE C. albicans BAX-associated cDNA fragment SEQ ID 447.

XX KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;

XX KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;

XX KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;

XX KW neurodegeneration; cell death; ss.

XX OS Candida albicans.

XX PN WO200264766-A2.

XX PD 22-AUG-2002.

XX PF 21-DEC-2001; 2001WO-EP015398.

XX PF 22-DEC-2000; 2000EP-00870318.

XX PR 04-JAN-2001; 2001EP-00870002.

XX PR 09-JAN-2001; 2001EP-00870003.

XX XX (JANC ) JANSSEN PHARM NV.

XX XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;

XX XX WPI; 2002-667002/71.

XX DR P-ESDB; ABG93245.

XX XX New isolated nucleic acid representing a synthetic BAX-gene, useful as

XX PT medicament for treating, preventing and/or alleviating yeast or fungal

XX PT infections or proliferative disorders, or for preventing apoptosis in

XX PT certain diseases.

PS Claim 36; Fig 2; 34app; English.

XX This invention describes a novel nucleic acid representing a synthetic

CC Bax gene. The Bax gene of the invention is useful for identifying Bax-

CC resistant yeast or fungi, identifying, or obtaining and identifying

CC Candida spp. sequences that are differentially expressed in a pathway

CC eventually leading to programmed cell death or identifying inhibitors or

CC inhibitor sequences of Bax-induced cell death. The products of the

CC invention have cytostatic, fungicide, immunosuppressive, virucide and

CC vasotropic activity and can be used in vaccines or for gene therapy. The

CC isolated nucleic acids, polypeptides, pharmaceutical compositions,

CC antisense molecules and antibodies are useful as medicaments or in

CC preparing a medicament for treating, preventing and/or alleviating

CC diseases associated with yeast or fungi or proliferative disorders, such

CC as cancer, or for preventing apoptosis in certain diseases. The compounds

CC or polypeptides, or the genetically modified organism are useful for

CC preparing a medicament for modifying the endogenous flora of humans and

CC other mammals. The vaccine is useful for immunising against yeast or

CC fungal infections. Apoptosis-related diseases include autoimmune disease,

CC ischaemia, diseases related with viral infections or neurodegenerations.

CC This sequence represents a polynucleotide associated with the Bax gene

CC described in the disclosure of the invention

XX SQ Sequence 1287 BP; 490 A; 237 C; 253 G; 307 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.0079	Length:	1287
Score:	141.00	Matches:	76
Percent Similarity:	39.24%	Conservative:	68
Best Local Similarity:	20.71%	Mismatches:	137
Query Match:	6.52%	Indels:	86
DB:	6	Gaps:	16

US-09-813-329-6 (1-409) x ABQ76511 (1-1287)

QY 36 GlnLeuIleProLeuVal---LeuGlyPheIleGlyLeuGlyLeuValAlaIleLeu 54

Db 196 CAACTACTACCAATAGCAAGATGAGTTCCTCAATACTCAAGATTTAGTTTGTATATT 255

QY 55 AlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeu----- 71

Db 256 AATGATTATGTTTCCAGAAATGAAGATTTGTCAAAGTTGAAGAGGATTTATCGAAATTC 315

QY 72 -----LysSerLeuLysArgValValAspAsnLeu 81

Db 316 TTACGACGCAAGAAATTTACCAAAAGTTTCTTAACAGTTGGAATCCATTTATGTAAGTG 375

QY 82 GlnGlnArg-----LeuGlyIleAsnTyrLeuAspGluPheAspGluPheGln 97

Db 376 GAAATCAAGAAAAAGAAAGCAACCAAGAACTCATCATCTGATGTAAGACTTCTCA 435

QY 98 LysGluTyrGluAsnAlaLeuIleAspTyrProLysLys-----ValAspGlyLeu 114

Db 436 TCTGAGAGTGAAAGTCTCACTTCGGACAGCAAGAAAGTCTCTCCTCAGATAGCAGCTCT 495

QY 115 ThrAspGluGluAspAspAspGlyLeuAspSerIleAlaAspAspGluAsp 134

Db 496 TCCTCAGACAGTGAAGTTCTTCTCAGACAGTGAAGTTCTTCATCAGACAGTGAAGAC 555

QY 135 AspAspValSerTyrSerSerValAsp---AspValGlyAlaAspTyrGluAspTyrThr 153

Db 556 AGCGATGACGAGGAAGACAAAGGAAGCAAGGAAGCAAGAAAGATACAAAGACAGCGAA 615

QY 154 AspMetLeuAsn---LysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThr 172

Db 616 GACAGCGAAAAACGAAAAAGTGGGAAGAGACAAAGACACACAGCTCTGATTCAAGTTCC 675

QY 173 ThrAlaGluGlyGluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspVal 192

Db 676 AGTTCCGACTCAAAATCTGATTGATCTCAGACTCAGACTCAGCTCCAGCTCTGATTCAAGTTCT 735

QY 193 PheAspAspPheThrSerTyrAsnAlaHisLysLysGlnGluArgLysSerArgSer 212













PA (UYOR-) UNIV OREGON HEALTH SCI.  
 XX Zonana J, Ferguson BM, Headon D, Overbeek P;  
 XX WPI: 2002-360478/39.  
 DR P-PSDB; ABB09082.  
 XX Novel purified hypohidrotic ectodermal dysplasias protein, useful for  
 PT promoting the development of hair follicles and tooth growth.  
 XX Example 1; Col 63-66; 86pp; English.  
 XX The present invention describes a purified hypohidrotic ectodermal  
 CC dysplasias protein (I) that promotes the development of hair follicles.  
 CC (I) is a promoter of hair follicle growth and tooth growth. (I) is useful  
 CC for the development of hair follicles. (I) is also useful for commercial  
 CC and clinical diagnostic testing for ectodermal dysplasia (ED). (I) is  
 CC also useful as a stimulant for hair and tooth growth, for stimulating  
 CC endocrine sweat gland development in individuals from whom the normal  
 CC sweating mechanism is compromised by disease or surgery, for stimulating  
 CC the growth of mammary epithelial tissue, either for reconstructing or  
 CC cosmetic purposes, and for promoting or maintaining differentiation of  
 CC breast epithelium. (I) is also useful for functional analysis, antibody  
 CC production and patient therapy. ABL51009 to ABL51122 and ABB09082 to  
 CC ABB09090 represent sequences used in the exemplification of the present  
 CC invention  
 XX SQ Sequence 1574 BP; 333 A; 497 C; 445 G; 299 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0.0186 Length: 1574  
 Score: 138.00 Matches: 96  
 Percent Similarity: 35.92% Conservative: 57  
 Best Local Similarity: 22.54% Mismatches: 147  
 Query Match: 6.38% Indels: 127  
 DB: 6 Gaps: 18

US-09-813-329-6 (1-409) x ABL51009 (1-1574)

QY 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLe 54  
 DB 349 AGGGAACAGCTGCTGCTCTTCTGGGTTCTTGGCTCTCGTGGCCCTCCACCTGCT 408  
 QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74  
 DB 409 GACGTTGGCTGCTAC-----CTAGAGTTGGCTCGA 441  
 QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94  
 DB 442 GTTGGCGGGAACGCTGGAGCGGAGTCCCGCTTGGC----- 478  
 QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114  
 DB 478 ----- 478  
 QY 114 uThrAspGluAspAspAspGlyAspGly-----LeuAs 127  
 DB 479 -----GGCTGGGACCCCTGGCACCTCTGGCCCTTAAG 513  
 QY 127 pSerIleAlaAsp---AspGluAspAspValSerTyrSerSerValAspValG 146  
 DB 514 CAGCCTCGTGGCTCGACCCCTGACGCCCATCACC----- 550  
 QY 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165  
 DB 551 -----AGTCACTTGGGCA 564  
 QY 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluThrAspSerAlaSerSe 185  
 DB 565 GCCGTCACTAAGCAGCAGCCATTGGAAACCGGAGAGCCGACCTCCACTCTGACTCCA 624  
 QY 185 rAlaSerAsnAsp-----AspAsnValPheAspPheThrSerTyrAsnAl 201  
 DB 185 -----

Db 625 GGACGGCACCAGATGGCCCTATTGAATTTCTTCTCCCTGATGAAAGCCACTACTCTGA 684  
 QY 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215  
 Db 685 AGAAGAAGATGGCGTGTTCGCCCAATAAAGAAAGCAAAAGCAATAGAGGAGATGG 744  
 QY 216 ---ValArgAsnGluGlnGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234  
 Db 745 CCCAGTTAAATAACAA-GAAAAAGGAAAGAAAGAGGACCTCTCTGGACCCCAATGGCCCTC 803  
 QY 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMetHi 254  
 Db 804 CAGACCCCCCAGGACCTCCAGGACCCCGAGGAAATTCAGGGATTCCTGGAA 863  
 QY 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270  
 Db 864 TTCAGGAACACTGTTATTGGGACCACCTGCTCTCCAGGTCCTCTCGTCTCTCAAGGAC 923  
 QY 271 -----ArgProAl 273  
 Db 924 CCCCTGGCTCCAGGACCTTCTGCTGCTGATAAAGCTGGAACTCGAGAAACCCAGC 983  
 QY 273 a-AlaHisPheHisLeuSerSerArgArgArgHisGlnGlySer---MetGlyTyrHisG 292  
 Db 984 CAGCTGTGTGCATCTACAGGC-----CAAGGCTCAGCAATTCAGTCAAGA 1031  
 QY 292 LyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307  
 Db 1032 ATGATCTTTTCAAGTGGAGTGTCTCAATGCTGCTCTCGATCACTATGAACCCCAAGGTGT 1091  
 QY 307 LyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValT 327  
 Db 1092 TTAAGCTACATCCCGCAGCGGGAGCTGGAGGTACTGGTGGACGCGCACCTTACTTCT 1151  
 QY 327 YrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345  
 Db 1152 ATAGTCAGGTAGAAGTATACATCACTTCACTGACTTTGCCAGCTAGAGTGGTGG 1211  
 QY 345 LngLyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365  
 Db 1212 TGGATGAGAGCCCTTCTCGAGTGCACAGCAGCATCTGAGACCGGCGGCAAGCCACAC--T 1268  
 QY 365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385  
 Db 1269 ACAACACTTGTATACCGCAGCGGTCTGCTCTCAAGGCCCGCGCAGAGATCGCCGTCA 1328  
 QY 385 ySAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405  
 Db 1329 AGATGGTGACGCTGACATCTCC-----ATCAACATGACGAAGCACACACCGTCTTTG 1382  
 QY 405 LyIlePheLysVal 409  
 Db 1383 GGGCCATCAGGCTG 1396

RESULT 15  
 ACD07895  
 ID ACD07895 standard; cDNA; 1574 BP.  
 XX AC ACD07895;  
 XX 15-AUG-2003 (first entry)  
 XX cDNA encoding human EDAL-II.  
 XX Human; ss, gene; EDAL-II; hair follicle development; tooth development;  
 KW sweat gland development; ectodermal disease; alopecia; skin trauma; burn;  
 KW x-linked hypohidrotic ectodermal dysplasia; XLHED; breast cancer; HED;  
 KW autosomal recessive hypohidrotic ectodermal dysplasia; hirsutism;  
 XX ectopic teeth.  
 OS Homo sapiens.  
 XX US2003023991-A1.  
 PN

XX PD 30-JAN-2003.  
 XX PF 04-DEC-2000; 2000US-00729658.  
 XX PR 09-JUL-1998; 98US-0092279P.  
 PR 15-DEC-1998; 98US-0112366P.  
 PR 29-JUN-1999; 99US-00342681.  
 XX (ZONA/) ZONANA J.  
 PA (FERG/) FERGUSON B M.  
 PA (HEAD/) HEADON D.  
 PA (OVER/) OVERBEEK P.  
 XX Zonana J, Ferguson BM, Headon D, Overbeek P;  
 XX WPI; 2003-456312/43.  
 DR P-PSDB; ABO01928.  
 XX  
 PT Increasing or decreasing hair follicle development, tooth development, or  
 PT sweat gland development, in a tissue, involves modulating ectodermal  
 PT dysplasia protein isoform II activity in the tissue.  
 XX Claim 13; Page 46-47; 102pp; English.  
 XX  
 CC The invention relates to a method of increasing or decreasing one or more  
 CC of hair follicle development, tooth development, or sweat gland  
 CC development, in a tissue, involves altering ectodermal dysplasia (EDAL)  
 CC isoform II (EDAL-II) activity in the tissue. The method is useful for  
 CC increasing or decreasing hair follicle development, tooth development or  
 CC sweat gland development, in a tissue of a patient suffering from an  
 CC ectodermal disease such as X-linked hypohidrotic ectodermal dysplasia  
 CC (XLHED), autosomal recessive hypohidrotic ectodermal dysplasia (HED) or  
 CC alopecia. The method is useful for decreasing hair follicle development  
 CC in a subject suffering from hirsutism. The method is also useful for  
 CC treating breast cancer and ectopic teeth, burns and trauma of skin due to  
 CC surgery. The present sequence represents cDNA encoding human EDAL-II  
 XX  
 SQ Sequence 1574 BP; 333 A; 497 C; 445 G; 299 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 0.0186 Length: 1574  
 Score: 138.00 Matches: 96  
 Percent Similarity: 35.92% Conservative: 57  
 Best Local Similarity: 22.54% Mismatches: 147  
 Query Match: 6.38% Indels: 127  
 DB: 8 Gaps: 18  
 US-09-813-329-6 (1-409) x ACD07895 (1-1574)  
 QY 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuValValAlaIleLe 54  
 DB 349 AGGGAACAGCTGCTCTCTCTGGGTCTTTGGGCTCTGCTGGCCCTCCACTGCT 408  
 QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGlyLeuIle 74  
 DB 409 GACCTTGCTGCTCTAC-----CTAGAGTGGCTCGGA 441  
 QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94  
 DB 442 GTTCGGCGGGAACGTGGAGCCGAGTCCGCCCTTGGC----- 478  
 QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114  
 DB 478 ----- 478  
 QY 114 uThrAspGluAspAspAspGlyAspGly-----LeuAs 127  
 DB 479 -----GGCTCGGCAACCCCTGGCACCCTCTGGCACCCTAAG 513  
 QY 127 pSerIleAlaAsp---AspGluAspAspAspValSerTyrSerValAspAspValG 146  
 DB 514 CAGCTCGGTGGCTCGACCTCGACCCCTGACAGCCCATCACC----- 550

QY 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThrGly-- 165  
 DB 551 -----AGTCACCTTGGCA 564  
 QY 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185  
 DB 565 GCGGTACCTTAAGCAGCAGCATTTGGAAACCGGGAAGAGCGGCGACTCCACTCTGACTCCCA 624  
 QY 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201  
 DB 625 GGACGGGCACCATGTCCTATTGATTTCTTCTTCCTGATGAAGAACCATCTACTCTGA 684  
 QY 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215  
 DB 685 AGAAGAAAGTAGGCGTGTTCGCGCAATAAAGAAAGCAATGAAGAGCAGATGG 744  
 QY 216 ---ValArgAsnGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234  
 DB 745 CCAGTTTAAACAA-GAAAGGGGAAAGAGCAGGACCTCTCTGGACCAATGGCCCTC 803  
 QY 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMethi 254  
 DB 804 CAGGACCCCGCAGGACCTCCAGGACCCCGCAGGAATTCAGGGATTCTCTGGAA 863  
 QY 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer-- 270  
 DB 864 TTCAGGAACAATGTTATGGGACCACTGGTCTCCAGGTCCTCTCTGGTCTCAAGGAC 923  
 QY 271 -----ArgProAl 273  
 DB 924 CCCTGGCCCTCCAGGACCTTCTGTGTGTGTATAGCTGGAACCTCGAAGTCCAGAAAACCCAGC 983  
 QY 273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292  
 DB 984 CAGCTGTGTGTGCTTACAGGGC-----CAAGGGTCAGCAATTCAGTCAAGA 1031  
 QY 292 LyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307  
 DB 1032 ATGATCTTTTCAGGTGGAGTGCTCAATGACTGTGCTCGCATCACTATGAACCCCAAGGTGT 1091  
 QY 307 LyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrValT 327  
 DB 1092 TTAAGCTACATCCCGCAGCGGGAGCTGGAGTACTGTGGACGCGCACTTCTTCACTCT 1151  
 QY 327 yTAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345  
 DB 1152 ATAGTCAGGTAGAAGTATATACATCACTTCACTGACTTTGCCAGCTATGAGGTGGTGG 1211  
 QY 345 lngLyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365  
 DB 1212 TGGATGAGAGCCCTTCTCTGAGTGCACACGAGCATCGACGCGGGAAGACCAAC---T 1268  
 QY 365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385  
 DB 1269 ACAACACTTGCTATACCGAGCGGTCTCCCTCCTCAAGCGCGGCGAGAGATCGCGCA 1328  
 QY 385 ysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405  
 DB 1329 AGATGTCGCGCTGACATCTCC-----ATCAACATGAGCAAGCACACCGTCTTTTG 1382  
 QY 405 lyIlePheLysVal 409  
 DB 1383 GGGCCATCAGGCTG 1396

Search completed: September 15, 2004, 08:23:37  
 Job time : 657 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 06:59:38 ; Search time 103 Seconds  
(without alignments)  
2203.640 Million cell updates/sec

Title: US-09-813-329-6

Perfect score: 2162

Sequence: 1 MTATLKPFFITPTSANDGF.....DRNAVLREGNNRSYFGIFKV 409

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delopt 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEV=xlp  
-Q/cgn2\_1/USPRO-spool\_p/US09813329/runat\_13092004\_102753\_21701/app\_query.fasta\_1.583  
-DB=Issued Patents NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09813329 -CGCN 1 1 128 @runat\_13092004\_102753\_21701 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	6.4	1176	4	US-09-342-681C-14
2	138	6.4	1574	4	US-09-342-681C-1
3	137	6.3	3034	4	US-09-976-594-636
4	133	6.2	3489	2	US-08-728-323A-1
5	133	6.2	3489	2	US-09-298-568-1
6	133	6.2	3489	4	US-09-410-399-1
7	133	6.2	3207	2	US-08-770-379-20
8	133	6.2	32207	3	US-08-757-669A-20
9	133	6.2	32207	4	US-09-230-371A-20
10	129	6.0	4599	1	US-08-431-080-27
11	129	6.0	4599	2	US-08-938-534-27
12	129	6.0	4599	4	US-09-345-294-27
13	138	6.4	1176	4	Sequence 14, Appl
14	138	6.4	1574	4	Sequence 1, Appl
15	137	6.3	3034	4	Sequence 636, App
16	133	6.2	3489	2	Sequence 1, Appl
17	133	6.2	3489	2	Sequence 1, Appl
18	133	6.2	3489	4	Sequence 20, Appl
19	133	6.2	3207	2	Sequence 20, Appl
20	133	6.2	32207	3	Sequence 20, Appl
21	129	6.0	4599	1	Sequence 27, Appl
22	129	6.0	4599	2	Sequence 27, Appl
23	129	6.0	4599	4	Sequence 27, Appl
24	138	6.4	1176	4	Sequence 39, Appl
25	138	6.4	1574	4	Sequence 15, Appl
26	137	6.3	3034	4	Sequence 3, Appl
27	133	6.2	3489	2	Sequence 3, Appl
28	133	6.2	3489	2	Sequence 3, Appl
29	133	6.2	3489	4	Sequence 3, Appl
30	133	6.2	3207	2	Sequence 3, Appl
31	133	6.2	32207	3	Sequence 3, Appl
32	129	6.0	4599	1	Sequence 5, Appl
33	129	6.0	4599	2	Sequence 5, Appl
34	129	6.0	4599	4	Sequence 5, Appl
35	138	6.4	1176	4	Sequence 39, Appl
36	138	6.4	1574	4	Sequence 15, Appl
37	137	6.3	3034	4	Sequence 3, Appl
38	133	6.2	3489	2	Sequence 3, Appl
39	133	6.2	3489	2	Sequence 3, Appl
40	133	6.2	3489	4	Sequence 3, Appl
41	133	6.2	3207	2	Sequence 3, Appl
42	133	6.2	32207	3	Sequence 3, Appl
43	129	6.0	4599	1	Sequence 5, Appl
44	129	6.0	4599	2	Sequence 5, Appl
45	129	6.0	4599	4	Sequence 5, Appl

13	128.5	5.9	4522	4	US-08-956-171E-39	Sequence 39, Appl
14	123	5.7	1422	4	US-09-485-077A-15	Sequence 15, Appl
15	121.5	5.6	1424	4	US-09-206-576-1	Sequence 1, Appl
16	120	5.6	4507	2	US-08-588-459A-3	Sequence 3, Appl
17	120	5.6	4507	2	US-08-487-826B-3	Sequence 3, Appl
18	120	5.6	4507	2	US-09-210-288-3	Sequence 3, Appl
19	119	5.5	2854	4	US-08-971-089-3	Sequence 3, Appl
20	118	5.5	1493	6	5340934-5	Patent No. 5340934
21	117.5	5.4	866	4	US-09-589-287B-37	Sequence 37, Appl
22	117.5	5.4	866	4	US-09-588-947A-37	Sequence 37, Appl
23	117.5	5.4	866	4	US-09-589-286A-37	Sequence 37, Appl
24	116	5.4	1236	3	US-09-105-343A-1	Sequence 1, Appl
25	116	5.4	2017	1	US-07-667-276A-3	Sequence 3, Appl
26	115.5	5.3	1478	4	US-09-216-393B-7	Sequence 7, Appl
27	115	5.3	30549	4	US-09-134-001C-322	Sequence 322, App
28	114.5	5.3	564	4	US-09-396-937-7	Sequence 7, Appl
29	114.5	5.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
30	114.5	5.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl
31	114	5.3	1512	4	US-09-252-991A-8290	Sequence 8290, Ap
32	114	5.3	2637	4	US-09-540-236-1481	Sequence 1481, Ap
33	113	5.2	956	1	US-08-431-080-25	Sequence 25, Appl
34	113	5.2	956	1	US-08-431-080-32	Sequence 32, Appl
35	113	5.2	956	2	US-08-938-534-25	Sequence 25, Appl
36	113	5.2	956	2	US-08-938-534-32	Sequence 32, Appl
37	113	5.2	956	4	US-09-345-294-25	Sequence 25, Appl
38	113	5.2	956	4	US-09-345-294-32	Sequence 32, Appl
39	113	5.2	49617	4	US-09-596-002-28	Sequence 28, Appl
40	112.5	5.2	1952	4	US-09-533-029-89	Sequence 89, Appl
41	112.5	5.2	2384	1	US-07-814-964-10	Sequence 10, Appl
42	112.5	5.2	2384	1	US-08-258-442-10	Sequence 10, Appl
43	112.5	5.2	2384	1	US-08-328-809-5	Sequence 5, Appl
44	112.5	5.2	2384	4	US-08-866-840-5	Sequence 5, Appl
45	112.5	5.2	2384	5	PCT-US92-11107-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-342-681C-14  
; Sequence 14, Application US/09342681C  
; Patent No. 6355782  
; GENERAL INFORMATION:  
; APPLICANT: Zonana et al.  
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins  
; FILE REFERENCE: 52978  
; CURRENT APPLICATION NUMBER: US/09/342,681C  
; CURRENT FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 60/092,279  
; PRIOR FILING DATE: 1998-07-09  
; PRIOR APPLICATION NUMBER: 60/112,366  
; PRIOR FILING DATE: 1998-12-15  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1176  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1176)  
US-09-342-681C-14

Alignment Scores:  
Pred. No.: 2.12e-05 Length: 1176  
Score: 138.00 Matches: 96  
Percent Similarity: 35.92% Conservative: 57  
Best Local Similarity: 22.54% Mismatches: 147  
Query Match: 6.38% Indels: 127  
DB: 4 Gaps: 18

US-09-813-329-6 (1-409) x US-09-342-681C-14 (1-1176)

Qy 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuGlyValValAlaIleLe 54

```

Db      108 AGGACACAGCTGCCTCTTCTCGGTTCTTTCGCTCTCGTGGCCCTCCACCTGCT 167
Qy      54 uAlaLeuThrIleTrpGlnThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
Db      168 GACGTTGTGCTCTAC-----CTAGAGTTGGCTCGGA 200
Qy      74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94
Db      201 GTTGGCGGGAGACCTGGAGCGAGTCCGCTTGGC----- 237
Qy      94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114
Db      237 ----- 237
Qy      114 uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127
Db      238 -----GGCTCGGCACCCCTGGCACCTCTGGCACCTTAAG 272
Qy      127 pSerIleAlaAsp---AspGluAspAspValSerTyrSerSerValAspAspValG1 146
Db      273 CAGCTCGGTGGCTCGACCTGACGCCCATCACC----- 309
Qy      146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165
Db      310 -----AGTCACCTTGGGCA 323
Qy      166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185
Db      324 GCCGTCACTTAAGCAGCAGCATTTGGAACCGGGAGAGCCGCACTCCCACTGACTCCCA 383
Qy      185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201
Db      384 GGACGGGCACAGATGGCCCTATTGAATTTCTTCTCTGATGAAAAGCCACTACTCGA 443
Qy      201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
Db      444 AGAAGAAAGTAGCGGTGTTCGCCGAATTAAGAAGCAAAAGCAATGAAGGAGCAGATGG 503
Qy      216 ---ValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234
Db      504 CCCAGTTTAAAAACAA-GAAAGAGGGAAGAAAGCAGGACCTCTCGGACCAATGGCCCTC 562
Qy      234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgArgMetHi 254
Db      563 CAGGACCCCGAGGACCTCCAGGACCCCGAGGACCCCGAGGAATTCAGGGATTCTCTGGAA 622
Qy      254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270
Db      623 TTCCAGGAACAACATGTTATTGGGACCACTGCTGCTCCAGGTCTCTCTGCTCAAGGAC 682
Qy      271 -----ArgProAl 273
Db      683 CCCCTGGCTCCAGGACCTTCTGGTCTGCTGATTAAGCTGGAATCTCGAGAAACACGAG 742
Qy      273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292
Db      743 CAGCTGTGGTGCATCTACAGGC-----CAAGGGTCAGCAATCAAGTCAAGA 790
Qy      292 lyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307
Db      791 ATGATCTTTTCAGGTGGAGTGCTCAATGACTGGTCTCGCATCACTATGAACCCCAAGGTG 850
Qy      307 lyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrValTr 327
Db      851 TTAAGCTTACATCCCGCAGCGGGAGCTGGAGGTACTGTGGACGACGACTTCACTCT 910
Qy      327 yrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345
Db      911 ATAGTCAGGTAGAGTATACATCAACTTCACTGACTTTCAGCTATGAGTGGTGG 970
Qy      345 lngLysThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365

```

```

Db      971 TGGATGAGAACCCCTTCTCTCAGTCACACGACGATCGAGACGGCAGACCAAC--T 1027
Qy      365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385
Db      1028 ACAACACTTGTATACCGCAGGCGTCTGCTCTCAAGGCCCGCAGAGATCGCGTCA 1087
Qy      385 yAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405
Db      1088 AGATGGTCGACGCTGACATCTCC-----ATCAACATGACCAAGCACACCCGCTTCTTTG 1141
Qy      405 lyIlePheLysVal 409
Db      1142 GGCCCATCAGGCTG 1155

RESULT 2
US-09-342-681C-1
; Sequence 1, Application US/09342681C
; Patent No. 6355782
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 52978
; CURRENT APPLICATION NUMBER: US/09/342,681C
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (242)..(1417)
US-09-342-681C-1

Alignment Scores:
Pred. No.: 3,34e-05 Length: 1574
Score: 138.00 Matches: 96
Percent Similarity: 35.92% Conservative: 57
Best Local Similarity: 22.54% Mismatches: 147
Query Match: 6.38% Indels: 127
DB: 4 Gaps: 18

US-09-813-329-6 (1-409) x US-09-342-681C-1 (1-1574)
Qy      35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuValValAlaIleLe 54
Db      349 AGGGAACAGCTGCCGTGCTCTCTCTGGGTTCTTTCGCTCTCGCTGGCCCTCCACCTGCT 408
Qy      54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
Db      409 GACGTTGTGCTGTAC-----CTAGAGTTGGCTCGGA 441
Qy      74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94
Db      442 GTTGGCGGGGAAACGTGGAGCCGAGTCCCGCTTGGC----- 478
Qy      94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114
Db      478 ----- 478
Qy      114 uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127
Db      479 -----GGTCGGGCAACCCCTGGCACCTCTGGCACCTTAAG 513
Qy      127 pSerIleAlaAsp---AspGluAspAspValSerTyrSerSerValAspAspValG1 146
Db      514 CAGCTCGGTGGCTCGACCTCGACCCCATCACC----- 550

```

```
Qy 146 yAlaaspTyrGluaspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165
Db 551 -----AGTCACTTGGGCA 564
Qy 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluThrAspSerAlaSerSe 185
Db 565 GCGTCACTTAAGCAGCAGCATTTGGAAACCGGGAGAGCCGACCTCCACTCTGACTCCCA 624
Qy 185 rAlaSerAsnAep-----AspAsnValPheAspAepPheThrSerTyrAsnAl 201
Db 625 GGACGGGCACCATGATGCTTATGAATTTCTTCTCCGTATGAAGAGCCATCTCTGA 684
Qy 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
Db 685 AGAAGAAAGTAGGCGTGTTCGCCGCAATAAAGCAAGCAATGATGAAGCAGCAGATGG 744
Qy 216 ---ValArgAsnGluGlnGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlySe 234
Db 745 CCCAGTTTAAAGCAA-GAAAGAGGAAAGAGAGCAGGACCTCCTGGAGCCCAATGGCCCTC 803
Qy 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMetHi 254
Db 804 CAGGACCCCGAGGACCTCAGGACCCCGAGGACCCCGAGGATTCAGGGATTCCTGGAA 863
Qy 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270
Db 864 TTCAGGAACAACATGTTATGGGACCACTGCTGCTCCAGGTCTCTGCTCAAGGAC 923
Qy 271 -----ArgProAl 273
Db 924 CCCTGGCCTCCAGGACCTTCTGTGCTGCTGATAAGCTGGAACCTCGAAGAACCCAGC 983
Qy 273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292
Db 984 CAGCTGGTGCATCTACAGGC-----CAGGCTCAGCAATTCAGTCAAGA 1031
Qy 292 lyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307
Db 1032 ATGATCTTTCAGGTGGAGTGTCAATGACTGCTGCTGCATCACTATGAACCCCAAGTGT 1091
Qy 307 lyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValT 327
Db 1092 TTAAGCTATACCTCCCGCAGCGGGAGCTGGAGGTACTGGTGGCGGACCTTACTTCACT 1151
Qy 327 yAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345
Db 1152 ATAGTCAGTAGAGTATACATCACTCACTGACTTTGCCAGCTATGAGTGGTGG 1211
Qy 345 lnglyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365
Db 1212 TGGATGAGAGCCCTTCTGTCAGTGCACAGCAGCATCGAGACGGGCAAGACCAAC---T 1268
Qy 365 alHisThrCysHisThrSerClyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385
Db 1269 ACAACATCTGTATACCGCAGCGCTGCTGCCCTCAAGCGCCCGCAGAGATCCCGCTCA 1328
Qy 385 yAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405
Db 1329 AGATGGTGACGCTGACATCTCC-----ATCAACATGAGCAGCACCACCGTTCITTG 1382
Qy 405 lyIlePheLysVal 409
Db 1383 GGGCCATCAGGCTG 1396
```

## RESULT 3

```
US-09-976-594-636
; Sequence 636, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
```

```
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 636
; LENGTH: 3034
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 238877.11
US-09-976-594-636
```

## Alignment Scores:

```
Pred. No.: 0.000119 Length: 3034
Score: 137.00 Matches: 68
Percent Similarity: 36.56% Conservative: 49
Best Local Similarity: 21.25% Mismatches: 107
Query Match: 6.34% Indels: 96
DB: 4 Gaps: 12
```

US-09-813-329-6 (1-409) x US-09-976-594-636 (1-3034)

```
Qy 20 PheProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuIlePro 39
Db 332 TTTCCTAGAGAACCACTGTAATCGATGTCGCGGAGAAACAACACTGAGCCCAAGATG 391
Qy 40 LeuValLeuGlyPheIleGlyLeuGlyLeuValValAla-----IleLeu 54
Db 392 ATGAAGCTGGGCGCACTGAATAGGAAGACACTTCGACAGAAAGAGCCGAGGCTATTT 451
Qy 55 AlaLeuThrIleTyrGlnThrArgValSerHisLeuAspLysGluLeuLysSerLeu 74
Db 452 AGTCTAATGACTGGCAATGTAACACTTCGACGAATGTGAATTTGGGCCAGAGATCAGAG 511
Qy 75 LysArgValValAsp-----AsnLeuGlnGlnArgLeuGly----- 86
Db 512 TGTAATATGTGTAATACTCCAAAGTATGCTAAATTAGAAGAAAGACAGGATATGGTGGT 571
Qy 87 -----IleAsnTyrLeuAsp-----GluPhe 93
Db 572 GGTTTTAAATGAAGAGAAAAATGTTGAATATATAGAAAGAGAAGAAATCTGATGGTGAATAT 631
Qy 94 AspGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGly 113
Db 632 GATGAGTTTGGCGTAAAAAGAAAAA-----TACAGAGGGAAGCAGTTTGGT 679
Qy 114 -----LeuThrAspGluGluAspAspAspGlyAspGlyLeuAspSerIleAla 130
Db 680 CTTGCATCTATATAAGGAAGTTGAAGATAAAGAAATCAGAGGGAGAGAA----- 730
Qy 131 AspAspGluAspAspValSerTyrSerValAspAspValGlyAlaAspTyrGlu 150
Db 731 GAGGATGAGGATCAAGATCTTCTAAATATAAGTAGAT----- 769
Qy 151 AspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGlyThrThrProThrSer 170
Db 769 ----- 769
Qy 171 GluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSerAlaSerAsnAsp 190
Db 770 -----GAGGATGAGGATGAAGATGAGCTGATCTCTCAAAATATAT- 811
Qy 191 AsnValPheAspAspPheThrSerTyrAsnAlaHisLysLysGlnGluArgLysSer 210
Db 812 -----CTTCATGCCAGTGAAGAGAGAGATAGTAATAAAAAAGAAATCTAATAGACGAAGT 865
Qy 211 ArgSerIleAlaAspValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeu 230
Db 866 CGTCA----- 871
```

QY 231 GlnGluLysSerSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArg 250  
 Db 872 AAGTCTCGATCTTCACATTCAGGATCTTCATCAGCTCATCTCTCCCTCAAGTTCAGG 931  
 QY 251 ArgArgMetHisSerArgHisHisLeuLeuValArgLysAlaArgSer----- 267  
 Db 932 TCTAGGTCAGGTCCTCGGTTCAAGAGTTCCTTCCAGTTCGCAGTCAAGATCTCGTTCCAGT 991  
 QY 268 -----GluAspSerArgProAlaHisPheHisLeuSerSerArgArgHisGln 285  
 Db 992 TCCAGAGAACGTTCCAGATCTCGTGGGTCGAAATCAAGATCCAGTCCAGTCCACAGG 1051  
 QY 286 GlySerMetGlyTyHisGlyAspMetTyTrileGlyAsnAspAsn-----GluArgAsn 303  
 Db 1052 GGCCTCTCTCCACGAAAGATCTTATTCAGTTTCATCTCTCTCGAGAGAAC 1111

## RESULT 4

US-08-728-323A-1  
 ; Sequence 1, Application US/08728323A  
 ; Patent No. 5948676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chang, Yuan  
 ; APPLICANT: Bohenzky, Roy A.  
 ; APPLICANT: Russo, James J.  
 ; APPLICANT: Edelman, Isidore S.  
 ; APPLICANT: Moore, Patrick S.  
 ; TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
 ; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA  
 ; TITLE OF INVENTION: Encoding Same And Uses Thereof  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/728,323A  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-278-0400  
 ; TELEFAX: 212-391-0525  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3489 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..3489  
 US-08-728-323A-1

Alignment Scores:  
 Pred. No.: 0.000405 Length: 3489  
 Score: 133.00 Matches: 59  
 Percent Similarity: 35.47% Conservative: 24  
 Best Local Similarity: 25.21% Mismatches: 108  
 Query Match: 6.15% Indels: 43  
 DB: 2 Gaps: 7

US-09-813-329-6 (1-409) x US-08-728-323A-1 (1-3489)  
 QY 91 AspGluPheAspGluPheGlnLysGluTyTrileGluAsnAlaLeuLeaSerTyProLysLys 110  
 Db 970 GATGACAATGACAATAAGGATGATGAGGAGGACGAGGACACATGAGGAGGACGAGGAG 1029  
 QY 111 ValAspGlyLeuThrAspGluAspAspAspAspGlyAspGlyLeuAspSerIleAla 130  
 Db 1030 GATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAG 1089  
 QY 131 AspAspGluAspAspValSerTyTrileGluAspValGlyAlaAspTyTrileGlu 150  
 Db 1090 GATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAG 1149  
 QY 151 AspTyThrAspMetLeuAsnLysLeuAsnAlaHisThrGlyThrThrProThrSer 170  
 Db 1150 GAT-----GACGAG 1158  
 QY 171 GluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSerAlaSerAsnAsp 190  
 Db 1159 GAGGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1218  
 QY 191 AsnValPheAspAspPheThrSerTyTrileGluAsnAlaHisLysLysLysGlnGluArgLysSer 210  
 Db 1219 -----GATGATGAGGACATGAGGAGGAGGATGACGAGGAGGAGGAGGAGGAGGAGGAG 1269  
 QY 211 ArgSerIleAlaAspValArgAsnGluGlnAsnIleGlnGlyAsnHisThr----- 228  
 Db 1270 -----GAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1311  
 QY 229 -----GluLeuGlnGlu-LysSer-----SerAsnGluAlaTh 239  
 Db 1312 ATCCAAAGTTCACACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1371  
 QY 239 rSerLysGluSerProAlaProLeuHis-----HisArgArgMetHisSe 255  
 Db 1372 CAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1431  
 QY 255 rArgHisArgHisLeuLeuValArgLysAlaArgSerGluAspSerArgProAlaHis 275  
 Db 1432 CAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1491  
 QY 275 sPheHisLeuSerSerArgArgArgHisGlnGlySerMetGlyTyTrileGlyAspMetTy 295  
 Db 1492 GAGCCAC---AACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1548  
 QY 295 rIleGlyAsnAspAsnGluArgAsnSerTyTrileGlyHis 308  
 Db 1549 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1588

## RESULT 5

US-09-298-568-1  
 ; Sequence 1, Application US/09298568  
 ; Patent No. 6322792  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kieff, Elliott D.  
 ; APPLICANT: Ballestas, Mary E.  
 ; APPLICANT: Kave, Kenneth M.  
 ; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
 ; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE  
 ; FILE REFERENCE: 16412-10001R  
 ; CURRENT APPLICATION NUMBER: US/09/298,568  
 ; EARLIER FILING DATE: 1999-04-21  
 ; EARLIER APPLICATION NUMBER: US 60/109,422  
 ; EARLIER FILING DATE: 1998-11-19  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3489  
 ; TYPE: DNA  
 ; ORGANISM: Kaposi's sarcoma-associated herpesvirus

US-09-298-568-1











```
; APPLICATION NUMBER: SN 08/326,781
; FILING DATE: October 20, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4599 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-938-534-27

Alignment Scores:
Pred. No.: 0.0017 Length: 4599
Score: 129.00 Matches: 73
Percent Similarity: 32.97% Conservative: 48
Best Local Similarity: 19.89% Mismatches: 126
Query Match: 5.97% Indels: 120
DB: 2 Gaps: 12

US-09-813-329-6 (1-409) x US-08-938-534-27 (1-4599)

QY 69 LysGluLeuLysSerLeuLysArgValValAspAsnLeuGlnArgLeuGlyIleAsn 88
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1114 AAAAAAGTAACTTATACAAAGGCAATACAAATGACGAGGCGC----- 1161
QY 89 TyrLeuAspGluPheAspGluPheGln-----LysGluTyrGluAsnAla 103
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1162 -----ACAGAGTATCGATTACAGCTTAACAGCTTAACAGAGGAGAGTGAATAAGAA 1215
QY 104 LeuIleAspTyrProLysLysValAspGlyLeuThrAspGluAspAspAspGly 123
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1216 GAAGAAGAGAGTGAAGAAGAGAGATGATGACGAGAGCAGCAGCAGATGATGAT 1275
QY 124 AspGluLeuAspSerIleAlaAspAspGluAspAspValSerSerValAsp 143
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1276 GACGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
QY 144 AspValGlyAlaAspTyrGluAsp----- 151
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1318 GACGAGATATAGATTGTGCAACTGACGCCCAAGAAAGAAAGAGCCATGAAGCT 1377
QY 152 -----TyrThrAspMetLeuAsnLysLeuAsn 160
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1378 TTATCTGCCATGAATACAAATAGCAATACACTCTATTCTCTCGTGAAGACAGTAACAAA 1437
QY 161 AsnAlaHisThrGlyThrProThrSerGluThrThrAlaGlu----- 175
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1438 AATAAATCAGTTAACTATCCCTAAGAAAGAAATGAAGAAGACGAGAAAGAAAGAAA 1497
QY 176 -----Gly 176
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1498 GAAAAAGAAAGAGACACAAAAACACAGAAATCAACAAAAAGAGTAACCGT 1557
QY 177 GluGlyLeuThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspPhe 196
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1558 TCAGGCCACTACTACTACACACAGCGCTATCGTTTAAATTCAAAAAGAGGACGACGCGC 1617
QY 197 ThrSerTyr-----AsnAlaHisLysLysGlnGlySerSerSerIleAlaAsp 215
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1618 ATTAGTTTGGTAATGGTAATGAAGGCTATTAACGAGGATATAGGTGAAGAAGTCTTGAT 1677
QY 216 ValArgAsnGluGluGlnAsn-----IleGln 224
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1678 TTAATAAACAAAGAGACAAATGGTAATGAAGAGATAAAGTGGATTCTAAGGTGATTTA 1737
QY 225 GlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAlaThrSerLysGluSerPro 244
```

```
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1738 GGTAACAACGATGAGTTACGATTTCCCAATATTTCAGAGTCAGATGAATCTGAA----- 1791
QY 245 AlaProLeuHisHisArgArgArgMetHisSerArgHisLeuLeuValArgLys 264
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1791 ----- 1791
QY 265 AlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArgHis 284
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1792 -----TATGATATTGACGAGGATCGCTACTTTGACGTGATTAAATGAAGATTCT 1842
QY 285 GlnGlySerMetGlyTyr-----HisGlyAspMetTyrIleGlyAsnAsp 299
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1843 CATGGAGAAATGGTACAGATCTTGAACCGGGGAGACGATCTTCCCATATTGGAAGAA 1902
QY 300 AsnGluArgAsn-----SerTyrGlnGly--- 307
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1903 GAAGAACAACAAATGTTTCTGAGCTACAAATACGACGAACTCTCATTCGATGGTAGT 1962
QY 308 ---HisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrVal 326
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1963 ATACACGAAGAGGCTGATCCTGTAGAAGATCCTGAAATAAATTTTGCAGAAATGAA 2022
QY 327 TyrAlaGln---IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheGln 345
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2023 TACAATCAGAAACCGATATGATGAAGAAGATGACCAAGAGATGAATAATGCTCTGAT 2082
QY 346 GlyAspThrProPheLeuGln 352
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2083 TTTGATATGCCGTTTATGAA 2103

RESULT 12
US-09-345-294-27
; Sequence 27, Application US/09345294
; Patent No. 6387619
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; SINGER, Miriam S.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/345,294
; FILING DATE: 30-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,080
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4599 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
i
US-09-345-294-27
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
Alignment Scores:
Pred. No.: 0.0017 Length: 4599
Score: 129.00 Matches: 73
Percent Similarity: 32.97% Conservative: 48
Best Local Similarity: 19.83% Mismatches: 126
Query Match: 5.97% Indels: 120
DB: 4 Gaps: 12

US-09-813-329-6 (1-409) x US-09-345-294-27 (1-4599)
QY 69 LysGluLeuLysSerLeuLysArgValValAspAsnGlnArgLeuGlyIleAsn 88
DB 1114 AAAAAAGTAACTTATCAAGGCGAATAGACAAATGACGAGGCGC-----1161
QY 89 TyrLeuAspGluPheAspGluPheGln-----LysGluTyrGluAsnAla 103
DB 1162 -----ACAGAGTCATCCGATTACCAAGCTGTAAACAGATGCGAAGAGAGTGAATAAGAA 1215
QY 104 LeuIleAspTyrProLysLysValAspGlyLeuThrAspGluGluAspAspAspGly 123
DB 1216 GAAGAAGAGAGTGAAGAAGAAGAAAGATGATGACGAAGACGACGACGATGATGAT 1275
QY 124 AspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSerSerValAsp 143
DB 1276 GACGGAAGTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
QY 144 AspValGlyAlaAspTyrGluAsp-----151
DB 1318 GACGAGATATAGATTTTGTCAAACTGACGCCCAAGAAAGAAAGACGATGAAAGCT 1377
QY 152 -----TyrThrAspMetLeuAsnLysLeuAsn 160
DB 1378 TTATCTGCCATGATACAAATAGCAATACACTTATTCCTCTCGTGAGAACAGTAAACAA 1437
QY 161 AsnAlaHisThrGlyThrProThrSerGluThrThrAlaGlu-----175
DB 1438 AATAAATCAGTTAAATATATCCCTTAAGAAAGAAATGAAGAGAGCAGAGAAGAAAGAA 1497
QY 176 -----Gly 176
DB 1498 GAAAAAGAGAAGAGAGCAACAAAAACAAGAAATCAAAACAAAAAAGAAAGTAAACGGT 1557
QY 177 GluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPhe 196
DB 1558 TCAGGCCTACTACTACACACAGCGCTATCGTTTAAATTCAAAAAGAGGACGACGGC 1617
QY 197 ThrSerTyr---AsnAlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAsp 215
DB 1618 ATTAGTTTGGTAATGTAATGAAGGCTATAACGAGGATATAGGTGAAGAGTCTTGGAT 1677
QY 216 ValArgAsnGluGluGlnAsn-----IleGln 224
DB 1678 TTAATAAACAAAGAGAAACAATGTTAATCAAGAGATAAACTGGATTCTAAGGTGATGTTA 1737
QY 225 GlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAlaThrSerLysGluSerPro 244
DB 1738 GGTAAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1791
QY 245 AlaProLeuHisHisArgArgArgMetHisSerArgHisLysLeuValArgLys 264
DB 1791 -----1791
QY 265 AlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArgHis 284
DB 1792 -----TATGATATGACAGGATGCGTACTTTGACGTGATTAACTAATGAAGATTC 1842
QY 285 GlnGlySerMetGlyTyr-----HisGlyAspMetTyrIleGlyAsnAsp 299
DB 1843 CATGGAGAAATGGTACAGATCTTGAAACGGGGAGACGATCTTCCCATATTTGGAGAA 1902
```

```

300 AsnGluArgAsn-----SerTyrGlnGly---- 307
1903 GAAGAAACAAACATTTCTGAGCTACAAAATGACGACGAACTCTCATTCGATGCTAGT 1962
308 ---HisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrVal 326
1963 ATACACGAGAGGGTCTGATCTCTGTAGAGATGCTGAAATTAATTTTTCGAAATGAA 2022
327 TyrAlaGln---IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheGln 345
2023 TACAATCAAGAAACGATATGAGGAAGATGACGAAGAAGATGAAATATGCTGAT 2082
346 GlyAspThrProPheLeuGln 352
2083 TTTGATATGCCGTTTATGAA 2103

RESULT 13
US-08-956-171E-39
; Sequence 39, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-08-956-171E-39

Alignment Scores:
Pred. No.: 0.00187 Length: 4522
Score: 128.50 Matches: 73
Percent Similarity: 30.25% Conservative: 48
Best Local Similarity: 18.25% Mismatches: 120
Query Match: 5.94% Indels: 159
DB: 4 Gaps: 13
```

```
US-09-813-329-6 (1-409) x US-08-956-171E-39 (1-4522)
Qy 60 GlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAsp 79
Db 2194 CAACAAAGCGTGAAGAACATAAAACGAAGCAAGATGCAATACAAACGTTGCAATTGAT 2253
Qy 80 AsnLeuGlnArgLeuGlyLeuAsnTyrLeu
Db 2254 GAAATGTATGCTAAACACGGAACGCTTATGTTGTTGATGATGTTCTATTAAATGATCATAGT 2313
Qy 90
Db 2314 GACTTACAGATAATAGTACAGATGCTAGTCAGCTTTCATACAAATGGCATAGAGAATGAA 2373
Qy 91
Db 2374 ACTGTATCAATGATGAATAATAACAAACGGCTCAATACAAATGAAGACACTAAATGACACT 2433
Qy 93 ---PheAspGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysVal 111
Db 2434 CATGTAGATGAAGTCCATACAAATATATGAGGAAGTTAGTTTGAATCAAGTATCGACAACA 2493
Qy 112 AspGlyLeuThrAspGluGlu
Db 2494 AAACAATTTGTCAGATGATGAAGTTACGGTTTCGAATGTAAAGCTCTCAACATCAATCAGCA 2553
Qy 119
Db 2554 CTACACATAACGTTGAATGAATGAATAAGATGAACATAAAATCAATCCAGATTAAT 2613
Qy 130 AlaAspAspGluAspAsp
Db 2614 GCTGATTGAGGAAGATGGACCAAGCAAGATAAGAGAGATATTCAGKAAGTCAATCGAT 2673
Qy 144 AspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHis 163
Db 2674 GAT
Qy 164 ThrGlyThrThrProThrSerGluThrAlaGluGlyGluGlyGluThrAsp
Db 2701
Qy 182
Db 2755 CGCTGTAATGCATGTAGACGCTCTCTAAACGCAAGAGTAGCAGTAACATCTCAAA 2814
Qy 186 AlaSerAsnAspAsnValPheAspAspPheThrSerTyrAsnAlaHisLysLysLys 205
Db 2815 GTAAATAATATCGATAAAACCGTTTGATAATGAAATTTGAATTTAGCACCGCGTCATAAAAAA 2874
Qy 206 GlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGluGlnAsn
Db 2875 GATGACCAACAAACTTAAAGTGTCAACTCTATGAAACGAATGATGTAATGATAATCAT 2934
Qy 223
Db 2935 GTTGTGGAGATTCAAGCATGAATCAATAGAAAGATAACGACGAGAAATTTACAGAAAT 2994
Qy 234 SerSerAsnGluAlaThrSerLysGlu
Db 2995 GTGCAAAACGAAGCAGCTGAAAGTGACAAATAATGTCGAAGAGAAAACTATTGAAAAACGTA 3054
Qy 243
Db 3055 AATCCAAAGAACACAGACTGAAAGGTTTCAACTTTAAGTAAAGACCACTTTAATGTTGTC 3114
Qy 249
Db 3115 ATGACGCCATCTGATAAAACCGTATGATGATCGTAAAGACATTCAAAAGTCAATGTG 3174
Qy 264
Db 3175 CCTGAATTAAGCCTGTACAAAGTAAGCAAGCTGTGAGTGAAAGATCCCTCGCAGTCAA 3234
Qy 276 PheHisLeuSerSerArgArgHisGlnGlySer
Db 3235 GCCACACCATCATCAGATCTGATTCAAGAGTCAAAATACAAATGATATATAAACAAAT 3294
Qy 294 MetTyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAsp 313
Db 3295 AATATGACATCAAAACAATGTTGAGAACAACTTATTGGTCTATGCAGAAACAGAAAAAT 3354

RESULT 14
US-09-485-077A-15
; Sequence 15, Application US/09485077A
; Patent No. 6458590
; GENERAL INFORMATION:
; APPLICANT: Mukherjee, Anil
; APPLICANT: Kundu, Gopal
; APPLICANT: Panda, Dibyendu
; TITLE OF INVENTION: Methods and Compositions for Treatment of Restenosis
; FILE REFERENCE: NIH-05047
; CURRENT APPLICATION NUMBER: US/09/485,077A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: PCT/US98/16569
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: 60/054,967
; PRIOR FILING DATE: 1997-07-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-485-077A-15

Alignment Scores:
Pred. No.: 0.00123 Length: 1422
Score: 123.00 Matches: 79
Percent Similarity: 33.14% Conservative: 38
Best Local Similarity: 22.38% Mismatches: 116
Query Match: 5.69% Indels: 120
DB: 4 Gaps: 17

US-09-813-329-6 (1-409) x US-09-485-077A-15 (1-1422)
Qy 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuIleProLeu 40
Db 19 CCAGTTGCAGCTTCTCAGCCAAACCCGACCAAGAAAACCTCTACCATGAGTAATGGCA 78
Qy 41 ValLeuGlyPheIleGlyLeuValValAlaIleLeuAlaLeuThrIleTrpGln 60
Db 79 GTGATTTGCTTTTGCCTCTAGGCATCACCTGTGCCATA-----CCAGTTAAACAG 129
Qy 61 ThrThrArgValSerHisLeuAspLysGluLeu----- 71
Db 130 GCTGATTTCTGAAGTCTGAGGAAAGACAGCTTTTACAAATAATACCCAGATGCTGTGCC 189
Qy 72 -----LysSerLeuLysArgValValAspAsnLeuGlnArgLeu 85
Db 190 ACATGGCTAAACCTGACCATCTCAGAACGAG-----AATCTCTAGCCCCACAG 240
Qy 86 GlyIleAsnTyrLeuAspGluPheGlnLysGluTyrGluAsnAlaLeuIle 105
Db 241 AATGCTGTGCTCTGAAGAACCAATGACTTTTAAACAACAG----- 282
Qy 106 AspTyrProLysLys-----ValAspGlyLeuThrAspGluGlu 118
Db 283 ACCCTTCCAAAGTAAGTCCAAACGAAAGCCATGACCACATGATGATGATGATGATGAT 342
Qy 119 AspAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspVal--- 137
Db 343 GATGATGACCATGTGGACACCCAGGACTCCATTGACTCGAACGACTCTGATGATGATGAT 402
Qy 138 -----SerTyrSerSerValAspValGlyAlaAspTyrGluAsp----- 151
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 10:23:37 ; Search time 744 Seconds  
(without alignments)  
2765.985 Million cell updates/sec

Title: US-09-813-329-6

Perfect score: 2162

Sequence: 1 MTATLKPFTPTTSANDGF.....DRNAVLREGNRSYFGIFKV 409

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09813329/runat\_13092004\_102755\_21784/app\_query.fasta\_1.583  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09813329 @CGN 1 1 912 @runat\_13092004\_102755\_21784  
-NCPU=6 -ICPU=3 -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*

14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*

17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*

18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*

19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

RESULT 1

US-09-813-329-5

; Sequence 5, Application US/09813329

; Patent No. US20020012968A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole.

; TITLE OF INVENTION: Variants Thereof

; FILE REFERENCE: D0016.np

; CURRENT APPLICATION NUMBER: US/09/813,329

; CURRENT FILING DATE: 2001-03-20

; PRIOR APPLICATION NUMBER: 60/190,816

; PRIOR FILING DATE: 2000-03-21

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 5

; LENGTH: 2148

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (634)..(1860)

US-09-813-329-5

ALIGNMENTS

						Sequence 5, Appli
						Sequence 3, Appli
						Sequence 1, Appli
						Sequence 45792, A
						Sequence 46034, A
						Sequence 447, Ap
						Sequence 52112, A
						Sequence 81649, A
						Sequence 7734, A
						Sequence 14, Appl
						Sequence 1, Appl
						Sequence 41, Appl
						Sequence 6, Appl
						Sequence 41, Appl
						Sequence 6, Appl
						Sequence 45, Appl
						Sequence 24791, A
						Sequence 19, Appl
						Sequence 19, Appl
						Sequence 271, Appl
						Sequence 19, Appl
						Sequence 1, Appl
						Sequence 107179, A
						Sequence 87071, A
						Sequence 69258, A
						Sequence 60017, A
						Sequence 58, Appl
						Sequence 39, Appl
						Sequence 39, Appl
						Sequence 61685, A
						Sequence 6421, Ap
						Sequence 8857, Ap
						Sequence 6025, A
						Sequence 22862, A
						Sequence 4557, A
						Sequence 2682, Ap
						Sequence 2682, Ap
						Sequence 16858, A
						Sequence 7853, Ap
						Sequence 95869, A
						Sequence 6546, A
						Sequence 13, Appl
						Sequence 230, Appl
						Sequence 175, Appl



Alignment Scores:  
Pred. No.: 1,08e-233 Length: 2148  
Score: 2162.00 Matches: 409  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-813-329-6 (1-409) x US-09-813-329-5 (1-2148)

QY 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe 20  
Db 634 ATGACTGCGGAGACCTCAAGCGTTTATAAGCCCAACGAGTGCACGATGATGTTT 693

QY 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40  
Db 694 CCGGCCAAGGACGACGACGCGCGACCGCCGAGCGACCGCGCGATGATCCCGCTG 753

QY 41 ValLeuGlyPheIleGlyLeuGluValValAlaIleLeuAlaLeuThrIleTrpGln 60  
Db 754 GTTTTGGGGTTTCATCGGTCTGGGGCTGGTGTGTCATTCGCACTAACGATCTGGCAG 813

QY 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn 80  
Db 814 ACAACGCGTGTATCGCATCTGCACAAAGGAGCTGAAGAGCCTGAAGCGAGTCTCGATAAT 873

QY 81 LeuGlnGlnArgLeuGlyLeuAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100  
Db 874 CTCACGACGCGTTGGGCATAACTATCTGACAGATTCGACGAGTTCGCAAGAGGATAC 933

QY 101 GluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluGluAspAsp 120  
Db 934 GAGAAATGCTCATCGACTATCCAAAAGGTGATGGCTCAGCGATGAGGAGGACGAC 993

QY 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer 140  
Db 994 GACGATGCGGATGCTGATTCATTCGACGACGACGACGACGACGATGATGATGATGATG 1053

QY 141 SerValAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn 160  
Db 1054 TCTGTGGATGATGTTGGCGAGACTACGAGGACTACACCGATATGTTAAATAAATCAAC 1113

QY 161 AsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180  
Db 1114 AATGCATATACCGCACCCACCATCTGAGACCATCTGCTGAGGGCGAGGAGCG 1173

QY 181 AspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyrAsn 200  
Db 1174 GACAGTGCATCTCAGCGCTCAAAATGATGACAAATGTTGTCGATGACTTTACGAGTCAAT 1233

QY 201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu 220  
Db 1234 GCCCACAANAAGACGAGGAGAGAAAATCTCGCTCGATTCGCCATGTACCAATAGGAG 1293

QY 221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAlaThrSer 240  
Db 1294 CAGAAATATCAAGAAATCACACAGAGCTTCAGAAAGTCAATCAATGAGGCAATCTCC 1353

QY 241 LysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHisLeu 260  
Db 1354 AAAGAGAGCCCTGCACCATCTCACCGCTCGAGAAATGCAATTCGCGCCATCGCCACCTC 1413

QY 261 LeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSer 280  
Db 1414 CTAGTCGCAAGCCAGATCCGAGGACTCGAGGCCAGAGCCCATTTCCATTTGAGCAGC 1473

QY 281 ArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn 300  
Db 1474 AGCGGCGTCAACAAAGAGATATGGCTACCATGGAGATATGATACATGAAATGATAAC 1533

QY 301 GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn 320

Db 1534 GAGAGAACTCTTATCAGGACACTTTTCAAAACGCGGATGGCTCTTGTGACGGTGACCAAT 1593

QY 321 ThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAsnGly 340  
Db 1594 ACAGGCTATATTACGTATAGCCCGAGATATGTACAACTCTCGACGACGACGAGCGGA 1653

QY 341 PheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsn 360  
Db 1654 TTTATCGTCTTTCAAGGAGACACTCCATTCCTCGTGTGTAACACGGTGGCCACCAAC 1713

QY 361 MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGlu 380  
Db 1714 ATGCCACATAGGTGACACCTGCCACACGAGTGGTCTGATCCACTTGAACGAAACGAG 1773

QY 381 ArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsn 400  
Db 1774 AGATCCATCTGAAGGACATTCACAAGATCGCAATGCAATTCGCGGGGAGGAAACAC 1833

QY 401 ArgSerTyrPheGlyIlePheLysVal 409  
Db 1834 CGAAGCTACTTTGGCATCTTCAAGGTG 1860

RESULT 2  
US-09-813-329-3  
; Sequence 3, Application US/09813329  
; Patent No. US20020012968A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole  
; TITLE OF INVENTION: Variants Thereof  
; FILE REFERENCE: D0016.np  
; CURRENT APPLICATION NUMBER: US/09/813,329  
; PRIOR FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,816  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: Patent version 3.0  
; SEQ ID NO 3  
; LENGTH: 1221  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1) .. (1218)  
US-09-813-329-3

Alignment Scores:  
Pred. No.: 4,44e-225 Length: 1221  
Score: 2082.50 Matches: 400  
Percent Similarity: 96.39% Conservativity: 0  
Best Local Similarity: 96.39% Mismatches: 0  
Query Match: 96.39% Indels: 15  
DB: 9 Gaps: 2

US-09-813-329-6 (1-409) x US-09-813-329-3 (1-1221)

QY 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe 20  
Db 1 ATGACTGCGGAGACCTCAAGCGTTTATAAGCCCAACGAGTGCACGATGATGTTT 60

QY 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40  
Db 61 CCGGCCAAGGACGACGACGCGCGACCGCCGAGCGACCGCCGAGCTGATCCCCCTG 120

QY 41 ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTrpGln 60  
Db 121 GTTTTGGGGTTTCATCGGTCTGGGGCTGGTGTGTCATTCGCACTTAACGATCTGGCAG 180

QY 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn 80  
Db 191 ACAACGCGTGTATCGCATCTGCACAAAGAGCTGAAGAGCTTGAAGAGGAGTCTCGATAAT 240

QY 81 LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100

```

Db      241  |||||||CTCCAGCAGCGCTTTGGGCATAAATCTATCTGACGAGTTTCGACGATTTCCAAAAGGAGTAC 300
Qy      101  GluAsnAlaLeuIleAaspTyrProLysValAaspGlyLeuThrAaspGluGluAasp 120
Db      301  GAGATGCGCCCTCATCGATCTCCAAAAGGTGATGCGCTTCAGGATGAGGAGGACGAC 360
Qy      121  AspAaspGlyAaspGlyLeuAaspSerIleAlaAaspAaspGluAaspAaspValSerTyrSer 140
Db      361  GACGATGCGGATGGTCTGATTCCATTTCGCGACGACGAGGACGACGAGCTTAGCTATAGC 420
Qy      141  SerValAaspAaspValGlyAlaAaspTyrGluAaspTyrThrAaspMetLeuAasnLysLeuAasn 160
Db      421  TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAATAAACTCAAC 480
Qy      161  AsnAlaHisThrGlyThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180
Db      481  AATGCACATACCGCACCCACCCACATCTGAGACCACTGCTGAGGGCGGGCGAGCG 540
Qy      181  AspSerAlaSerSerAlaSerAsnAaspAaspValPheAaspAaspPheThrSerTyrAsn 200
Db      541  GACAGTGCACTCTCAGCGCTCAAAATGATGACAATGTGTTTCGATGACTTTTACCAGCTACAAT 600
Qy      201  AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAaspValArgAasnGluGlu 220
Db      601  GCCCACAAAAAGAACGACGAGAGAAAATCTCGCTCGATTGCGGATGTACGCAATGAGGAG 660
Qy      221  GlnAsnIleGlnGlyAasnHisThrGluLeuGlnGluLysSerSerAasnGluAlaThrSer 240
Db      661  CAGAAATATTCAAGGAAATACACAGAGCTTCAGGAAAAGTCAATCAATGAGGCAACTTCC 720
Qy      241  LysGluSerProAlaProLeuHisThrHisArgArgMetHisSerArgHisArgHisLeu 260
Db      721  AAAGAG-----AGATGCATTCCCGCCATCGCCACCTC 753
Qy      261  LeuValArgLys-----AlaArgSerGluAaspSerArgProAlaAla 274
Db      754  CTAGTCCCAAGGTGAATCTCTTTTCAGCCACAGATCCGAGGACTCGAGGCCAGCAGCC 813
Qy      275  HisPheHisLeuSerSerArgArgHisGlnGlySerMetGlyTyrHisGlyAaspMet 294
Db      814  CATTTCCACTTTGAGCAGCGCGCGCTCACCAGGAAGTATGGGCTACCATGGGAGATG 873
Qy      295  TyrIleGlyAasnAaspAasnGluArgAasnSerTyrGlnGlyHisPheGlnThrArgAaspGly 314
Db      874  TACATAGGAATGATACGACGAGAACTCTTATCAGGACACTTTCAAAACGCGGATGGC 933
Qy      315  ValLeuThrValThrAasnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAasnAasn 334
Db      934  GTCTTGACGGTGACCAATACAGGCTATATTACGTATACGCCACGATATGCTACAACAAC 993
Qy      335  SerHisAaspGlnAasnGlyPheIleValPheGlnGlyAaspThrProPheLeuGlnCysLeu 354
Db      994  TCGCACGACCAACGACGAACTTATCGCTTTTCAAGGAGACATCTCCATTCCTGCACTGCTG 1053
Qy      355  AsnThrValProThrAasnMetProHisLysValHisThrCysHisThrSerGlyLeuIle 374
Db      1054  AACACGGTCCCAACCAATCCACATPAAGGTGCACACCTGCCACACGAGTGGTCTGATC 1113
Qy      375  HisLeuGluArgAasnGluArgIleHisLeuLysAaspIleHisAasnAaspArgAasnAlaVal 394
Db      1114  CACCTGGAACGAAACGAGAGGATCCATCTGAAGGACATTCACAACGATCGCAATGCGATT 1173
Qy      395  LeuArgGluGlyAasnAasnArgSerTyrPheGlyIlePheLysVal 409
Db      1174  CTGGGGAGGGAAACAACCGAAGTACTTTGGCATCTTCAAGGTTG 1218
```

RESULT 3

US-09-813-329-1  
; Sequence 1, Application US/09813329  
; Patent No. US20020012968A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company

```

; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole
; FILE OF INVENTION: Variants Thereof
; FILE REFERENCE: D0016.np
; CURRENT APPLICATION NUMBER: US/09/813,329
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2166
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (652)..(1878)
US-09-813-329-1
Alignment Scores:
Pred. No.: 1,84e-220 Length: 2166
Score: 2045.00 Matches: 394
Percent Similarity: 96.82% Conservative: 2
Best Local Similarity: 96.33% Mismatches: 13
Query Match: 94.59% Indels: 0
DB: 9 Gaps: 0
US-09-813-329-6 (1-409) x US-09-813-329-1 (1-2166)
Qy      1  MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAasnAaspGlyPhe 20
Db      652  ATGACTGCCGAGACCCCTCAAGCGCTTTATTAACCCCAACGAGTGCCCAACGATGATGGT 711
Qy      21  ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40
Db      712  CCGGCCAANACGACACGACGCGACCGCCGACGACGACCCGCGCAGCTGATCCCCCTG 771
Qy      41  ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTrpGln 60
Db      772  GTTTTGGGGTTTCATCGCTCTGGGCTGGTGGTTCGTCATTCTCGCATTACGATCTGGCAG 831
Qy      61  ThrThrArgValSerHisLeuAaspLysGluLeuLysSerLeuLysArgValValAaspAasn 80
Db      832  ACAACGCGTGTATCGCATCTCGGACAAAGGAGCTGAAGAGCCTGAAGCGAGTCTCGATAAT 891
Qy      81  LeuGlnGlnArgLeuGlyIleAasnTyrLeuAaspGluPheAaspGluPheGlnLysGluTyr 100
Db      892  CTCAGCAGCGCTTTGGGCATAAACTATCTCGACGAGTTTCGACGAGTTTCCAAAAGGAGTAC 951
Qy      101  GluAasnAlaLeuIleAaspTyrProLysValAaspGlyLeuThrAaspGluAaspAasp 120
Db      952  GAGAATGCCCTCATCGACTATCCAAAAAAGGTGGATGGCTCTACGGATGAGGAGGACGAC 1011
Qy      121  AspAaspGlyAaspGlyLeuAaspSerIleAlaAaspAaspGluAaspAaspValSerTyrSer 140
Db      1012  GACGATGGCGATGGTCTCGGATTCCTATTCGCGACGACGAGGACGACGCTTAGCTATAGC 1071
Qy      141  SerValAaspAaspValGlyAlaAaspTyrGluAaspTyrThrAaspMetLeuAasnLysLeuAasn 160
Db      1072  TCTGTGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAATAAACTCAAC 1131
Qy      161  AsnAlaHisThrGlyThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180
Db      1132  AATGCACATACCGCACCCACCCACATCTGAGACCACTCTGAGGCGGAGGGCGAGCG 1191
Qy      181  AspSerAlaSerSerAlaSerAasnAaspAasnValPheAaspAaspPheThrSerTyrAasn 200
Db      1192  GACAGTGCATCTCAGCCTCAAAATGATGACAATGTGTTTCGATGACTTTTACCAGCTCAGAT 1251
Qy      201  AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAaspValArgAasnGluGlu 220
Db      1252  GCCCTCAAAAAGAACGACGAGGAGAAAATCTCGCTCGAATTCGCCATGTACCAATGAGGAG 1311
Qy      221  GlnAasnIleGlnGlyAasnHisThrGluLeuGlnGluLysSerSerAasnGluAlaThrSer 240
```

Db 1312 CAGAAATATTCAAGGAATACACAGAGCTTCAGGAAAGTCAATCAATGAGCGAGCTTCC 1371  
Qy 241 LysGluSerProAlaProLeuHisHisArgArgMetHisSerArgHisArgHisLeu 260  
Db 1372 AAAGAGAGCCCTCGAGCACTTTCACCTCGTCGCGAATGCATTCGCCCATCGCACCTC 1431  
Qy 261 LeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSer 280  
Db 1432 GTAGTCCGCAAGCAGATCCGAGACTTCGAGGACGAGCGCCCATTTCCACTTGAGCAGC 1491  
Qy 281 ArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn 300  
Db 1492 AGCGCGCTACCAAGAAAGTATGGCTTACCATGGAGATATGTACATAGAAAATGATAGG 1551  
Qy 301 GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn 320  
Db 1552 GAGAGATGCTCTTATCAAGGACACTTTCAAACGGCGATGGCGTATTTGACGGTGACCAAT 1611  
Qy 321 ThrGlyLeuTyrTyrValTyrAlaGlnTleCysTyrAsnAsnSerHisAspGlnAsnGly 340  
Db 1612 GCAGGCTATATTAGTATAGCCGACGATATGGGGCTACAACTCGCACGCCAGAACGGA 1671  
Qy 341 PheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsn 360  
Db 1672 TTTATCGTCTTCAAGGAGACACTCCATTCCTCGAGTGTGAAACACGGTGCCCAACAC 1731  
Qy 361 MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGlu 380  
Db 1732 ATGCCACATAAGGTGCACACCTGCCACAGAGTGGTCTGATCCACCTGGAAACGAAACGAG 1791  
Qy 381 ArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsn 400  
Db 1792 AGGATCCATCTGAAGGACATTCACACGATCGCAATGCAGTTCCTCGGGAGGGAACAAC 1851  
Qy 401 ArgSerTyrPheGlyIlePheLysVal 409  
Db 1852 CGAAGTACTTTGGCATCTTCAAGGTG 1878

## RESULT 4

US-10-437-963-45792/c

; Sequence 45792, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 45792

; LENGTH: 1765

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_48722C.1

US-10-437-963-45792

Alignment Scores:

Pred. No.: 6.79e-06 Length: 1765

Score: 145.50 Matches: 58

Percent Similarity: 43.32% Conservative: 49

Best Local Similarity: 23.48% Mismatches: 85

Query Match: 6.73% Indels: 55

DB: 17 Gaps: 11

US-09-813-329-6 (1-409) x US-10-437-963-45792 (1-1765)  
Qy 56 LeuThrIleTyrPheGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLys 75  
Db 1411 CTCACCTACTTCGAGTACCGCGCGTCAAGAGA--GAGAAGCTCCGCAAGAGCATGAAG 1355  
Qy 76 ArgValValAspAsnLeu-----GlnGlnArgLeuGly-----Ile 87  
Db 1354 CGGTGCATCTGGAACATCACGCGCCAGGCTCTCGCCGGAAGGTGAGATGAGGATAT 1295  
Qy 88 AsnTyrLeuAspGluPheAspGluPheGlnLysGluAsnAlaLeuIleAspTyr 107  
Db 1294 GGTACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1250  
Qy 108 ProLysLysValAspGlyLeuThrAspGluGluAspAspAspAspGlyAspGlyLeuAsp 127  
Db 1249 CCAAGAAAGGTG---GCCTCCTCGGATAGAGCGAAGAGGAGGATAGCAAGGTTTCGTG 1193  
Qy 128 SerIleAlaAspAspGluAspAspValSerTyrSerValAspAspValGlyAla 147  
Db 1192 GAATCCGACTCCGGTGAGTCTGATAGCTGTGCTCGATTCCAGCAAAATCAGAT----- 1142  
Qy 148 AspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThrGlyThrThr 167  
Db 1141 -----GATACCCCGAGGAAG---AAGAAAGGGCGTAAGGGTAGTCAC 1103  
Qy 168 ProThrSerGluThr-----ThrAlaGluGlyGluGlyGlu 179  
Db 1102 CGTAGCAGCAAGCGTAGCCGCCGCCGCCGCCCATTCATCTGTATACAGAGGGTGCAC 1043  
Qy 180 ThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheThrSerTyr 199  
Db 1042 ---GACAACAGCAAGCTCGAGGAGTTCGGAGGCTCTTATGATTCGAGGACTCTATG 986  
Qy 200 AsnAlaHisLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGlu 219  
Db 985 GATAGGAGGAAGAAGAGATCGCGG-----AGGCACAAG 950  
Qy 220 GluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSerSerAsnGlu----- 237  
Db 949 AAGTCTTAAGAGAGGGGGAAGGAGCTCCAGAGGAAGAAGAGGAAGCAATGATACAGCT 890  
Qy 238 -----AlaThrSerLysGluSerProAlaPro 246  
Db 889 TCTGAGGAGAGCTCTGAGGAGGAGCGGTGGCAGCGGCTTCAGGATCAAGCCTTAGCCCC 830  
Qy 247 LeuHisHisArgArgMetHisSerArgHisArgHisLeuLeuValArgLysAlaArg 266  
Db 829 TTGAGGGACAGTAAGNAGAAAGAGCCGAGCTCACGGAGGAGGAGGAGGAGGAGGAGGAG 770  
Qy 267 SerGluAspSerArgProAla 273  
Db 769 TCGGAAGATCAAGCCCCATCC 749

## RESULT 5

US-10-424-599-46034

; Sequence 46034, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J

; APPLICANT: Kovalic, David K

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 46034

; LENGTH: 1269

; TYPE: DNA

[illegible]



## Alignment Scores:

Pred. No.: 1.93e-05 Length: 1057  
Score: 138.50 Matches: 55  
Percent Similarity: 43.10% Conservative: 45  
Best Local Similarity: 23.71% Mismatches: 81  
Query Match: 6.41% Indels: 51  
DB: 13 Gaps: 8

US-09-813-329-6 (1-409) x US-10-424-599-81649 (1-1057)

```
Qy 59 TrpGlnThrThrArg-ValSerHisLeuAspLysGluLeuLysSerLeuLysArgValVa 78
Db 184 TGGAAACGTAAACGCTCTGCTGCTTCTCGAAGATGAAGGCCACAAAGAAATTGAG 243
Qy 78 LaspAsnLeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheGlnLys 98
Db 244 CGAGCAAGTGGAGGAAAC---GAGGTCAAGGTTTGTGAGGAGGACGACGAGAATA 300
Qy 98 sGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluGl 118
Db 301 CGAGTCCGTGGATGAAGTAGTAGACGACGACAGAA-----GATGATGAAGACGAAGA 354
Qy 118 uaspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSe 138
Db 355 GGACGATCAGCAGCAGGAGGACGACAAAT-----GACGACGAGGAAGATGAT----- 403
Qy 138 rTyrSerSerValAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLy 158
Db 404 -----GCTCTGACGGCGGTGACGACGACGACGACGAAGATGAT----- 442
Qy 158 sLeuAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGl 178
Db 443 -----GAGGAAGAAGG 453
Qy 178 yGluThrAspSerAlaSerAlaSerAsnAspAspValPheAspAspPheThrSe 198
Db 454 TGATGTCCAGCGGTGCTGTGAGCCTGACGATGACGATACGACTCAGACGACGACTCAGA 513
Qy 198 rTyAsnAlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAs 218
Db 514 CGACGACCAATGAAGACGAGGAGGACGACGAGGAG-----CAAGG 555
Qy 218 nGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSerAsnGluAl 238
Db 556 GGAAGACGAGGATTG-----GGAACAGAGTACCTTATCCGCCCTTTAGAAACTGC 606
Qy 238 aThrSerLysGluSerProAlaPro-LeuHisHisArgArgArgMetHisSerArgHisA 258
Db 607 TGAGAGGAGGAAGCATCAAGTGATTTTGAACCAAGAGGAGGAATGCTGAGGAGGAGGAAGA 666
Qy 258 rGHisLeuLeuValArgLysAlaArgSerGluAspSerArgProAlaHisPheHisL 278
Db 667 AGATGTTCCAGCAGGAGGACTGTGAGA-----AGGC 696
Qy 278 euSerSerArgArgArgHisGlnGlySerMet 288
Db 697 TGAGCTCCACCGAAAGAGAGGTCAGATA 728
```

## RESULT 9

US-10-425-114-7734  
; Sequence 7734, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 7734  
; LENGTH: 966  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700727423\_FLI  
US-10-425-114-7734

Alignment Scores:  
Pred. No.: 1.91e-05 Length: 966  
Score: 138.00 Matches: 60  
Percent Similarity: 38.52% Conservative: 34  
Best Local Similarity: 24.59% Mismatches: 90  
Query Match: 6.38% Indels: 61  
DB: 13 Gaps: 7

US-09-813-329-6 (1-409) x US-10-425-114-7734 (1-966)

```
Qy 67 LeuAspLysGluLeuLysSerLeuLysArgValValAspAsnLeuGlnArgLeuGly 86
Db 88 CTGATTTGAAGATGAAGGCCACAAAGAAATTGAGCCAGCATGTGGAGGAAATGAGGTC 147
Qy 87 IleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyrGluAsnAlaLeuIleAsp 106
Db 148 ACGTCTCTGTTGATGAGGAGGAGGACGAAGATACCGATCCGAGGATGAAGTAGTAGAC 207
Qy 107 TyrProLysLysValAspGlyLeuThrAspGluAspAspGlyAspGlyLeu 126
Db 208 -----GACGGAGGAGGACGATGAAGACGACGACGACGACGATGACGATGAC 252
Qy 127 AspSerIleAlaAspAspGluAspAspValSerTyrSerSerValAspAspVal--- 145
Db 253 GACGACAAATGAAGACGAGGAGGACGACGAT-----GAAGATGATGCTCTCCC 297
Qy 146 ---GlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThr 164
Db 298 GACGCGGTGACGACGACGACGATGAAGAC----- 327
Qy 165 GlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSer 184
Db 328 -----GAAGAAGAAAGATGATGTCACGCGCGGTGGA 360
Qy 185 SerAlaSerAsnAspAspAsnValPheAspAspPheThrSerTyrAsnAlaHisLysLys 204
Db 361 GAGCTCAGCAGCAGCAGTAAACGATGACGACGAC----- 393
Qy 205 LysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlnAsnIleGln 224
Db 394 -----GAAGATCAACACGAGGAGGATGAGGAGGACAA 426
Qy 225 GlyAsnHisThrGluLeu-GlnGluLysSerSerAsnGluAlaThrSerLysGluSerPr 244
Db 427 GGGAGAGGCGGATTTGGGAACAGAGTACCTTATCCGCCCTCTAGTCACTG----- 478
Qy 244 oAlaProLeuHisHisArgArgMetHis-----SerArgHisArgHisLeuLeu 261
Db 479 -----CTGAGGAGGAGGAGCATCTAGTATTTCGAACCCAGAGGAGAAATGGC 525
Qy 261 uValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerAr 281
Db 526 GAGGAGGAGGAGGAGATGCTCGATGATGAGGATGGTGAAGTCTGAGGCTCCACCAAAA 585
Qy 281 gArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsnGl 301
Db 586 AGAAAGAGGTCAGATAAAGATGATTTCAGATGACGATGATGCGCGGTGAA-GATGATGATGA 644
Qy 301 uArgAsnSer 304
Db 645 GAGACCCTCT 654
```

```

RESULT 10
US-09-729-658B-14
; Sequence 14, Application US/09729658B
; Publication No. US20030023991A1
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729,658B
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1176)
US-09-729-658B-14
Alignment Scores:
Pred. No.: 2,58e-05 Length: 1176
Score: 138.00 Matches: 96
Percent Similarity: 35.92% Conservative: 57
Best Local Similarity: 22.54% Mismatches: 147
Query Match: 6.38% Indels: 127
DB: 10 Gaps: 18

US-09-813-329-6 (1-409) x US-09-729-658B-14 (1-1176)
QY 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuValValAlaIleLe 54
DB 108 AGGAACAGCTGCTGCTTCCTCGGGTTCCTTGGGCTCTCGTGGCCCTCCACCTGCT 167
QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
DB 168 GACGTGTGCTGCTAC-----CTAGAGTTGGCTCGGA 200
QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyLysLeuAsnTyrLeuAspGluPheAs 94
DB 201 GTTCGGCGGGAACGTCGGAGCCGAGTCCCGCTTGGC----- 237
QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuLeuAspTyrProLysLysValAspGlyLe 114
DB 237 ----- 237
QY 114 uThrAspGluAspAspAspGlyAspGly-----LeuAs 127
DB 238 -----GGTCGGGACCCCTGGCACCTCTGGCACCTAG 272
QY 127 pSerIleAlaAsp---AspGluAspAspValSerTyrSerValAspAspValG 146
DB 273 CAGCCTCGGTGGCTCGACCTGACGCCATCACC----- 309
QY 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165
DB 310 -----AGTCACCTTGGCA 323
QY 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185
DB 324 GCCGTACCTAAGCAGCAGCATTGGAAACCGGAGAGCGGCACTCCACTCTGACTCCCA 383
QY 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201
DB 384 GGACGGCACCAAGATGCCCTATTGAAATTTCTTCTCCCTGATGAAAGCAACCTACTCTGA 443

```

---

```

QY 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
DB 444 AGAAGAAAGTAGCGGTTCGCGCAATAAAAGAACCAATGAAGAGCAGATGG 503
QY 216 ---ValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234
DB 504 CCAGTTTAAACAA-GAAAAGGGAAAGAAAGCAGGACCTCTCGGACCAATGGCCCTC 562
QY 234 rSerAsnGluAlaThrSerLysGluSerProLeuHisHisHisHisArgArgMethi 254
DB 563 CAGGACCCCGCAGGACCTCCAGGACCCCGGACCCCGAGGAATTCAGGATTCCTGGAA 622
QY 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer-- 270
DB 623 TTCAGGAACAACCTGTTATGGGACCACTGGTCTCTCAGGTCTCTCTGCTCAAGGAC 682
QY 271 -----ArgProAl 273
DB 683 CCCCTGGCTCCAGGACCTTCTGGTGTGCTGATATAAGCTGGAACCTCGAGAAACCAAC 742
QY 273 a-AlaHisPheHisLeuSerSerArgArgArgHisGlnGlySer---MetGlyTyrHisG 292
DB 743 CAGCTGTGTCATCTACAGGCG-----CAAGGGTCAGCAATTCAAGTCAAGA 790
QY 292 LyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307
DB 791 ATGATCTTTCAGGTGGAGTGCTCAATGACTGTGTCATCACTATGAACCCCAAGGTGT 850
QY 307 LyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValT 327
DB 851 TTAAGCTACATCCCGCAGCGGGAGCTGGAGTACTGGTGACCGCCTACTTCACTCT 910
QY 327 YrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345
DB 911 ATAGTCAGGTAGAAGTATACTACATCACTTCCACTGACTTTCAGTATGAGTGGTGG 970
QY 345 LngLysPthrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysEv 365
DB 971 TGGATGAGAGCCCTTCTTCAGTGCACACGAGCATCGAGACGGGCAAGCAAC---T 1027
QY 365 aHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385
DB 1028 ACAACACTGTCTATACCGCAGCGCTGCTCTCTCAAGCCCGCGCAGGAAGATCGCGTCA 1087
QY 385 yAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405
DB 1088 AGATGTGCACGCTGACATCTCC-----ATCAACATGAGCAAGCACACCACTTCTTTG 1141
QY 405 LyIlePheLysVal 409
DB 1142 GGGCATCAGGCTG 1155

```

---

```

RESULT 11
US-09-729-658B-1
; Sequence 1, Application US/09729658B
; Publication No. US20030023991A1
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729,658B
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1574
; TYPE: DNA

```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (242)..(1417)
US-09-729-658B-1

Alignment Scores:
Pred. No.:      4,02e-05      Length:      1574
Score:          138.00      Matches:      96
Percent Similarity: 35.92%      Conservative: 57
Best Local Similarity: 22.54%      Mismatches: 147
Query Match:      6.38%      Indels:      127
DB:              10      Gaps:      18

US-09-813-329-6 (1-409) x US-09-729-658B-1 (1-1574)

QY 35 ArgGlnLeuLeuPro-LeuValLeuGlyPheLeGlyLeuGlyLeuValValAlaIleIe 54
DB 349 AGGGAACAGTCCTGCTCTCTCTGGGTTCTTTGGCTCTCGCTGGCCCTCCACCTGCT 408
QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
DB 409 GACGTTGTGCTGCTAC-----CTAGAGTTGGCTCGGA 441
QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94
DB 442 GTTGGCGGGGAACGTGGAGCGAGTCCGCCCTTGGC-----478
QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114
DB 478 -----478
QY 114 uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127
DB 479 -----GGTCGGGCACCCCTGGCACCTCTGGCACCTTGGCACCTTAAG 513
QY 127 pSerIleAlaAsp---AspGluAspAspValSerTyrSerSerValAspValGl 146
DB 514 CAGCCTCGTGGCTCGACCTCGACCCCATCAC-----550
QY 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThrGly-- 165
DB 551 -----AGTCACCTTGGGCA 564
QY 166 -ThrThrProThrSerGluThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185
DB 565 GCCGTCACTAAGCAGCAGCATTTGGAACCGGGAGACCCGACCTCACTCTGACTCCA 624
QY 185 rAlaSerAsnAsp-----AspAsnValPheAspPheThrSerTyrAsnAl 201
DB 625 GGACGGGCACCATGATGATGATTTCTTCTTCTGATGAAAGCCATACCTCTGA 684
QY 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
DB 685 AGAAGAAAGTAGGCGGTTCGCCCGCAATAAAGAGCAAGCAATGAAGAGCAGATGG 744
QY 216 ---ValArgAsnGluGlnGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlyLysSe 234
DB 745 CCCAGTTAAACAA-GAAAAGGGGAAGAACGAGGACCTCTCGGACCCCAATGGCCCTC 803
QY 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMetHi 254
DB 804 CAGGACCCCGAGGACCTCCAGACCCCGAGGACCCCGAGGAATTCAGGGATTCTCGAA 863
QY 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270
DB 864 TTCCAGGAACAACTGTTATGGGACCACCTGCTCTCCAGGTCCTCTGCTCTCAAGGAC 923
QY 271 -----ArgProAl 273
DB 924 CCCCTGGCTCCAGGACCTTCTGGTGTGCTGATAAAGCTGGAACTCGAAGAACACAGC 983
QY 273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292
```

```
Db 984 CAGCTGTGGTGCATCTACAGGGC-----CAAGGGTCAAGCAATTCAGTCAAGA 1031
QY 292 lyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307
DB 1032 ATGATCTTTTCAGGTGGAGTGTCTCAATGACTGCTGCTCGCATCATGATGAACCCCAAGGTGT 1091
QY 307 lyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValT 327
DB 1092 TTAAGCTACATCCCGCAGCGGGAGCTGGAGGTACTGGTGGAGCGCACCTACTTCTATCT 1151
QY 327 yrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345
DB 1152 ATAGTCAGGTAGAAGTATACTACATCACTTCACTGACTTTTCCAGCTATATGAGGTGGTGG 1211
QY 345 lNGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365
DB 1212 TGGATGAGAACCCCTTCTGCGATGCGACACGACGCGGCGGCAAGACCAAC---T 1268
QY 365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385
DB 1269 ACAACACTTGCTATACCGCAGCGCTCTGCTCTCAAGGCCCGCAGAGAAGATCGCCGTCA 1328
QY 385 ysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405
DB 1329 AGATGGTGCAGCTGACATCTCC-----ATCAACATGAGCAGCAGCACACACAGCTTCTTTG 1382
QY 405 yIlePheLysVal 409
DB 1383 GGGCCATCAGGCTG 1396

RESULT 12
US-10-202-062-41
; Sequence 41, Application US/10202062
; Publication No. US20040038349A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.,
; TITLE OF INVENTION: Heteromultimeric TNF Ligand Family members
; CURRENT APPLICATION NUMBER: US/10/202,062
; PRIOR FILING DATE: 2002-07-25
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 5307
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4242)..(4242)
; OTHER INFORMATION: n equals a, g, c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4471)..(4471)
; OTHER INFORMATION: n equals a, g, c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4492)..(4492)
; OTHER INFORMATION: y equals c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4523)..(4523)
; OTHER INFORMATION: n equals a, g, c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4529)..(4529)
; OTHER INFORMATION: n equals a, g, c or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4531)..(4531)
; OTHER INFORMATION: n equals a, g, c or t
```



```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4545)..(4545)
; OTHER INFORMATION: n equals a, g, c or t
US-10-202-062-41

Alignment Scores:
Pred. No.: 0.000255 Length: 5307
Score: 138.00 Matches: 96
Percent Similarity: 35.92% Conservative: 57
Best Local Similarity: 22.54% Mismatches: 147
Query Match: 6.38% Indels: 127
DB: 13 Gaps: 18

US-09-813-329-6 (1-409) x US-10-202-062-41 (1-5307)

Qy 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLe 54
Db 350 AGGAACAGCTGCTGCTCTCTCTGGGTCTTTTGGGCTCTCGTGGCCCTCCACCTGCT 409
Qy 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
Db 410 GACCTTGCTGCTGCTAC-----CTAGAGTTGGCTCGGA 442
Qy 74 uLysArgValValAspAsnLeuGlnArgLeuGlyLeuAsnTyrLeuAspGluPheAs 94
Db 443 GTTCGGGGGAGACGTGGAGCGGAGTCCGCCCTTGGC----- 479
Qy 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114
Db 479 ----- 479
Qy 114 uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127
Db 480 -----GGCTCGGGGACCCCTGGCACCTCTGGCCACCTTAAG 514
Qy 127 pSerIleAlaAsp---AspGluAspAspAspValSerTyrSerSerValAspValG1 146
Db 515 CAGCTCGTGGGCTGACCTGACAGCCCATCACC----- 551
Qy 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165
Db 552 -----AGTCACCTTGGCA 565
Qy 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185
Db 566 GCCGTCACTTAAGCAGCAGCATTTGGAAACCGGGAGAACCGCACTCCACTCTGACTCCCA 625
Qy 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201
Db 626 GGACGGGCACCATGATGCCCTATTGAATTTCTTCTCCCTGATGAAAGCCATACCTGA 685
Qy 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
Db 686 AGAAGAAAGTAGGCGTGTTCGCCCAATAAAGAAAGCAAAAGCAATGAAGGAGCAGATGG 745
Qy 216 ----ValArgAsnGluGlnGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlyLysSe 234
Db 746 CCCAGTTAAAAACAA--GAAAAAGGGAAGAAAGCAGGACCTCTCTGGACCCCAATGGCCCTC 804
Qy 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMetHi 254
Db 805 CAGGACCCCGAGGACCTCCAGGACCCCGAGGACCCCGAGGATTCAGGGATTCCTGGAA 864
Qy 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270
Db 865 TTCCAGGAACAACATGTTATGAGCACCATGCTGCTCTCCAGGTCTCTCTGCTCAAGGAC 924
Qy 271 -----ArgProAl 273
Db 925 CCCCTGGCTCCAGGACCTTCTGGTGTCTGATTAAGGTGNACTCGAATAAACCCAGC 984
Qy 273 a-AlaHisPheHisLeuSerSerArgArgArgHisGlnGlySer---MetGlyTyrHisG 292
```



```

; LOCATION: (4523)..(4523)
; OTHER INFORMATION: n equals a, c, g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4529)..(4529)
; OTHER INFORMATION: n equals a, c, g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4531)..(4531)
; OTHER INFORMATION: n equals a, c, g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4545)..(4545)
; OTHER INFORMATION: n equals a, c, g or t
; US-10-218-547-41

Alignment Scores:
Pred. No.: 0.000255 Length: 5307
Score: 138.00 Matches: 96
Percent Similarity: 35.92% Conservative: 57
Best Local Similarity: 22.54% Mismatches: 147
Query Match: 6.38% Indels: 127
DB: 15 Gaps: 18

US-09-813-329-6 (1-409) x US-10-218-547-41 (1-5307)

QY 35 AtcGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuValValAlaIleLe 54
Db 350 AGGGAACAGCTGCTCTCTCTGGGTTCTTTCGGCTCTCGTGGCCCTCCACCTGCT 409
QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
Db 410 GAGCTTGTGCTGCTAC-----CTAGAGTTGGCTCGGA 442
QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94
Db 443 GTTGGCGGGAGACGTGGAGCCGAGTCCGCTTGGC-----479
QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114
Db 479 -----479
QY 114 uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127
Db 480 -----GGTTCGGGCACCCCTGGCACCTCTGGCACCCCTAAG 514
QY 127 pSerIleAlaAsp---AspGluAspAspValSerTyrSerSerValAspAspValG 146
Db 515 CAGCCTCGGTGGCTCGACCTCGAGCCCATCAC-----551
QY 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThrGly-- 165
Db 552 -----AGTCACCTTGGCA 565
QY 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185
Db 566 GCGCTCACCTAAGACAGCAGCCATCGAACCCGGAGAGCCGACCTCCACTCTGACTCCCA 625
QY 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201
Db 626 GGACGGGCACCATAGTGGCCCTATTGAAATTTCTTCTTCCTGATGAAAAGCCATACCTGA 685
QY 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
Db 686 AGAAGAAAGTAGCGGTGTTCGCCGCAATAAAAGAAAGCAAAAGCAATGAAGGAGCAGATGG 745
QY 216 ----ValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGluLysSe 234
Db 746 CCCAGTTAAACAA-AAAAAGGAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804
QY 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMethI 254
Db 805 CAGGACCCCCCAGGACCTCCAGGACCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 864

```

; OTHER INFORMATION: n = a, t, c or g  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4529)..(4529)  
 ; OTHER INFORMATION: n = a, t, c or g  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4531)..(4531)  
 ; OTHER INFORMATION: n = a, t, c or g  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4545)..(4545)  
 ; OTHER INFORMATION: n = a, t, c or g  
 ; US-10-272-328A-6

## Alignment Scores:

Pred. No.: 0.000255 Length: 5307  
 Score: 138.00 Matches: 96  
 Percent Similarity: 35.92% Conservative: 57  
 Best Local Similarity: 22.54% Mismatches: 147  
 Query Match: 6.38% Indels: 127  
 DB: 15 Gaps: 18

US-09-813-329-6 (1-409) x US-10-272-328A-6 (1-5307)

```

Qy 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuGlyLeuValAlaIleLe 54
    |||||
Db 350 AGGGAACAGCTGCTGCTCTTCTCGGTTCTTTCGCTCTCGTGGCCCTCCACCTGCT 409
    |||||

Qy 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
    |||||
Db 410 GACGTTGTGCTGTAC-----CTAGAGTTGCGCTCGGA 442
    |||||

Qy 74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94
    |||||
Db 443 GTTGGCGGGGAACCTGGAGCGAGTCCCGCTTGGC----- 479
    |||||

Qy 94 pGluPheGlnLysGluIuTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114
    |||||
Db 479 ----- 479

Qy 114 uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127
    |||||
Db 480 -----GGCTCGGGCACCCCTCGCACCTCTGGCACCCCTAAG 514
    |||||

Qy 127 pSerIleAlaAsp---AspGluAspAspValSerTyrSerValAspValG1 146
    |||||
Db 515 CAGCCTCGGTGGCTCGACCCCTGACGCCCATCAC----- 551
    |||||

Qy 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165
    |||||
Db 552 -----AGTCACCTTGGCA 565
    |||||

Qy 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185
    |||||
Db 566 GCCGTCACTAAGCAGCAGCATTTGGAACCCGGAGAGCCGCACTCCCACTCTGACTCCA 625
    |||||

Qy 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201
    |||||
Db 626 GGACGGGCACAGATGGCCCTATTGAAATTCTTCTTCCTGATGAAAGCCATACCTCTGA 685
    |||||

Qy 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
    |||||
Db 686 AGAAGAAGTAGCGGTGTTCCCGCAATAAAGAAGCAAAAGCAATGAAGGAGCAGATGG 745
    |||||

Qy 216 ----ValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234
    |||||
Db 746 CCCAGTTAAACAA-GAAAAAGGGAAAGAGCAGGACCTCCTGGACCAATGGCCCTC 804
    |||||

Qy 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisArgArgArgMetHi 254
    |||||
Db 805 CAGGACCCCGGACCTCCAGGACCCCGAGGACCCCGAGGAATTCAGGGATTCTCTGGAA 864
    |||||
  
```

```

Qy 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270
    |||||
Db 865 TTCCAGGAACAACACTGTTATGGGACCACCTGGTCTCTCCAGGTCTCTCTGGTCTCAAGGAC 924
    |||||

Qy 271 -----ArgProAl 273
    |||||

Db 925 CCCCTGGCCTCCAGGGACCTTCTGGTCTGCTGATNAAGCTGGAACTCGAGAAACCCAGC 984
    |||||

Qy 273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292
    |||||
Db 985 CAGCTGTGGTCACTACAGGGC-----CAAGGGTCAGCAATTCAGTCAAGA 1032
    |||||

Qy 292 lyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307
    |||||
Db 1033 ATGATCTTTCAGGTGGAGTGTCAATGACTGCTCGCATCAGTATGAACCCCAAGGTGT 1092
    |||||

Qy 307 lyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValT 327
    |||||
Db 1093 TTAAGCTACATCCCGGAGGGGAGCTGGAGGTACTGGTGGACGGCACCTTACTTCTATCT 1152
    |||||

Qy 327 YrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345
    |||||
Db 1153 ATAGTCAGGTAGAAGTATACATCACTTCACTGACTTTGCCAGCTATGAGGTGGTGG 1212
    |||||

Qy 345 lnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365
    |||||
Db 1213 TGGATGAGAAGCCCTCTCTGCAGTGCACACGACGATCGAGACGGCAAGACCAAC---T 1269
    |||||

Qy 365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385
    |||||
Db 1270 ACAACACTTGTATACCGCAGGCGTCTGCTCTCAAGGCCCGGCAGAAAGATCCCGTCA 1329
    |||||

Qy 385 ysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405
    |||||
Db 1330 AGATGGTGACCGCTGACATCTCC-----ATCAACATGACGACACACACACGTTCTTTG 1383
    |||||

Qy 405 lyIlePheLysVal 409
    |||||
Db 1384 GGGCCATCAGGCTG 1397
    |||||
  
```

Search completed: September 15, 2004, 13:25:39  
 Job time : 763 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 05:12:07 ; Search time 4190 Seconds  
(without alignments)  
2914.948 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTATLTKPFTTPTSDNGF.....DRNAVLREGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US9813329/runat\_13092004\_102753\_21680/app\_query.fasta\_1.583  
-DB=EST -QFMT=fas+cap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US9813329 @CGN 1 1 6425 @runat\_13092004\_102753\_21680 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	843	39.0	686	12	BG636194 SD13795.5
2	829	38.3	495	12	B1636564
3	480	22.2	508	14	CK135985
C 4	466	21.6	532	14	CK135525
C 5	397	18.4	612	13	BX560278
6	206.5	9.6	539	13	BX560278
7	193.5	9.0	526	9	AI253821
8	172	8.0	1101	29	CNS017MQ
9	170	7.9	917	29	CNS07DXT
10	151	7.0	709	13	BU373036
11	150	6.9	965	14	CF590046
12	145	6.7	746	14	CD851068
13	143	6.6	550	12	B1514804
14	143	6.6	724	13	BU798281
C 15	143	6.6	732	13	BU803616
16	143	6.6	951	11	AY223179
17	143	6.6	1148	14	CD509148
18	143	6.6	2660	11	BC049115
19	141	6.5	746	13	BC620580
20	140.5	6.5	1413	11	BC020355
21	140.5	6.5	1414	11	BC029991
22	140.5	6.5	2894	11	BC023777
23	140	6.5	688	29	AY401339
24	140	6.5	703	13	BU445245
25	140	6.5	892	13	BU911456
26	138.5	6.4	291	13	BX560279
27	137.5	6.4	670	12	BX269011
28	137	6.3	906	12	B1250058
29	137	6.3	2812	11	BC033648
30	136	6.3	648	12	BM615338
31	135.5	6.3	800	13	BU219475
32	135	6.2	713	13	BU117025
33	135	6.2	936	11	CNS08SL1
34	134.5	6.2	582	10	BF482404
35	134.5	6.2	866	12	B1249241
36	134	6.2	814	12	B1250570
37	134	6.2	875	14	CA473339
38	134	6.2	914	12	B1250061
39	134	6.2	923	13	BU771970
40	133	6.2	615	14	CA855938
41	133	6.2	912	13	BU149401
C 42	132.5	6.1	580	12	BJ379142
C 43	132.5	6.1	679	12	BM203188
44	132.5	6.1	715	12	BJ617087
45	132.5	6.1	780	12	BI108540

ALIGNMENTS

RESULT 1  
BG636194 686 bp mRNA linear EST 23-APR-2001  
LOCUS SD13795.5prime SD Drosophila melanogaster Schneider L2 cell culture  
DEFINITION pot2 Drosophila melanogaster cDNA clone SD13795 5 similar to  
CG12919: FBan0012919 located on: 2R 46E1-46E1;: 04/13/2001, mRNA  
sequence.  
ACCESSION BG636194  
VERSION BG636194.1 GI:13763731  
KEYWORDS EST.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 686)  
**REFERENCE**  
**AUTHORS** Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,  
 Lewis,S. and Rubin,G.M.  
**TITLE** BDGP/HMI Drosophila EST Project  
**JOURNAL** Unpublished (2001)  
**COMMENT** Contact: Stapleton, M.  
 BDGP

Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
 hit genomic AB003831: arm:2R [4876906,5137815]  
 estimated-cyto:46B13-46E4: 04/13/2001  
 Plate: SD.137 row: H column: 11  
 High quality sequence stop: 685.  
 Location/Qualifiers

## FEATURES

1..686

/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="SD13795"  
/lab\_host="DH5-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
culture pOT2"  
/note="Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized  
fractionated cDNAs were directly ligated into pOT2.  
Plasmid cDNA library."

## ORIGIN

Alignment Scores:  
Pred. No.: 5,376-65 Length: 686  
Score: 843.00 Matches: 153  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.99% Indels: 0  
DB: 12 Gaps: 0  
US-09-813-329-6 (1-409) x BG636194 (1-686)

QY 257 HisArgHisLeuLeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPhe 276  
 Db 2 CATGCCACCTCTCTAGTCCGCAAGCCAGATCCGAGGACTCGAGGCCAGAGCCCATTC 61  
 QY 277 HisLeuSerSerArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIle 296  
 Db 62 CACCTGAGCAGCAGCGCGGCTCACCAAGGAAGTAGGGCTACCATGGAGATATGTACATA 121  
 QY 297 GlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeu 316  
 Db 122 GGAATGATAACGAGAGAAACTCTTATCAGGGACACTTTCAACGCCGCGATGGCTCTTG 181  
 QY 317 ThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHis 336  
 Db 182 ACGGTGACCAATACAGCCCTATATATAGGTATAGCCAGATATGCTACAAACATCGCAC 241  
 QY 337 AspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThr 356  
 Db 242 GACCAGAACGGATTATCGTCTTTCAGAGAGACACTCCATCTCTGCAAGTCTTGAACACG 301  
 QY 357 ValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeu 376  
 Db 302 GTGCCCAACCAATGCCACATAGGTGCACACCTGCCACACGAGTGTCTGATCCACCTG 361  
 QY 377 GluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArg 396  
 Db 362 GAACGAACGAGAGATCCCATCTGAGGACATTCACAAACATCGCAATGCAATGCTCTGCGG 421  
 QY 397 GluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409  
 Db 422 GAGGAAACCAACCAAGCTACTTTGGCATCTTCAAGGTG 460

## RESULT 2

Bi636564

## LOCUS

## DEFINITION

Bi636564 495 bp mRNA linear EST 10-SEP-2001  
 SD18286.5prime SD Drosophila melanogaster Schneider L2 cell culture  
 pOT2 Drosophila melanogaster cDNA clone SD18286 5 similar to  
 CG12919: F8an0012919 GO:() located on: 2R 46E1-46E1;: 05/19/2001,  
 mRNA sequence.

## ACCESSION

Bi636564

## VERSION

Bi636564.1 GI:15538774

## KEYWORDS

EST.

## SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 495)

Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,

Lewis,S. and Rubin,G.M.

BDGP/HMI Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)

hit genomic AB003831: arm:2R [4876906,5137815]

estimated-cyto:46B13-46E4: 05/19/2001

Plate: SD.182 row: H column: 2

High quality sequence stop: 491.

Location/Qualifiers

1..495

/organism="Drosophila melanogaster"

/mol\_type="mRNA"

/db\_xref="taxon:7227"

/clone="SD18286"

/lab\_host="DH5-alpha"

/clone\_lib="SD Drosophila melanogaster Schneider L2 cell

culture pOT2"

/note="Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized

fractionated cDNAs were directly ligated into pOT2.

Plasmid cDNA library."

Alignment Scores:

Pred. No.: 6,068-64 Length: 495

Score: 829.00 Matches: 162

Percent Similarity: 99.39% Conservative: 0

Best Local Similarity: 99.39% Mismatches: 1

Query Match: 38.34% Indels: 1

DB: 12 Gaps: 0

US-09-813-329-6 (1-409) x Bi636564 (1-495)

QY 100 TyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluAsp 119

Db 6 TACGAGATGCCTCATCGACTATCCCAA-AAGGTGGATGGCTCACGGATGAGGAGAC 64

QY 120 AspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyr 139

Db 65 GACGACGATGGCGATGGTCTGGATTCCATTGCGGACGAGGACGAGCGAGCTTAGCTAT 124

QY 140 SerSerValAspAspValGlyValAspTyrGluAspTyrThrAspMetLeuAsnLysLeu 159

Db 125 AGCTCTGTGGATGATGTTGGCGACACTACCGGACTACCGGATATGTTAAATAAATC 184

QY 160 AsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGlu 179

Db 185 AACATGCACATACCGGCACCCACCGCCCATCTGAGACCATCTGCTGAGGGCGAGGCGAG 244

QY 180 ThrAspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyr 199

Db 245 ACGGACGTGCTATCCTCAGCCTCAAAATGATGACAAATGTTGTCGATGACTTTTACCAGCTAC 304

Qy 200 AsnAlaHisLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGlu 219  
 |||||  
 Db 305 AATCCCAAAAAGAACGACGAGAGAAATCTCGCTCGATTGCCGATGTACGCAATGAG 364  
 |||||  
 Qy 220 GluGlnAenIleGlnGlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAlaThr 239  
 |||||  
 Db 365 GAGCAGATATATCAAGGAAATACACAGAGCTTCAGGAAAGTCAATCCCAATGAGGCAACT 424  
 |||||  
 Qy 240 SerLysGluSerProAlaProLeuHisHisArgArgMetHisSerArgHisArgHis 259  
 |||||  
 Db 425 TCCAAGAGAGCCCTGCACCACTTACCACCCTCGCAGATGCATTCCTCCGCCATCGCCAC 484  
 |||||  
 Qy 260 LeuLeuVal 262  
 |||||  
 Db 485 CTCCTAGTC 493

RESULT 3  
 CK135985/c  
 LOCUS  
 DEFINITION SD18286.3prime SD Drosophila melanogaster Schneider L2 cell culture  
 POT2 Drosophila melanogaster cDNA clone SD18286 3, mRNA sequence.

CK135985  
 CK135985.1 GI:38627921  
 EST.  
 Drosophila melanogaster (fruit fly)  
 ORGANISM  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
 1 (bases 1 to 508)  
 Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,  
 Lewis,S. and Rubin,G.M.

TITLE  
 JOURNAL  
 COMMENT BDGP/HMI Drosophila EST Project  
 Unpublished (2001)  
 Other ESTs: SD18286.5prime  
 Contact: Stapleton, M.

BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
 Based upon the presence of a XhoI site followed by a run of 14 or  
 more T residues at the beginning of the sequence, this clone was  
 polyadenylated. The resulting Poly-T sequence has been removed. hit  
 genomic AE003831: arm:2R [4876906,5137815]  
 estimated-cyto:46B13-46E4: 02/08/2002  
 Plate: SD.182 row: H column: 2  
 High quality sequence stop: 456.

FEATURES  
 source  
 1..508  
 Location/Qualifiers  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"  
 /clone="SD18286"  
 /lab\_host="DHS-alpha"  
 /clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
 culture POT2"  
 /note="vector: pOT2; Site\_1: EcoRI; Site\_2: XhoI; Sized  
 fractionated cDNAs were directly ligated into pOT2.  
 Plasmid cDNA library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 8.99e-33 Length: 508  
 Score: 480.00 Matches: 86  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 22.20% Indels: 0  
 DB: 14 Gaps: 0

US-09-813-329-6 (1-409) x CK135985 (1-508)

Qy 324 TyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisEspGlnAsnGlyPheIleVal 343  
 |||||  
 Db 508 TATTACGTATACGCCAGATATGCTACAACTCGCACGACGAAACGGATTATCGTC 449  
 |||||  
 Qy 344 PheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHis 363  
 |||||  
 Db 448 TTTCAAGGAGACACTCCATTCTCGCATGCTTGAACACGGTCCCAACATGCCACAT 389  
 |||||  
 Qy 364 LysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluAatgIleHis 383  
 |||||  
 Db 388 AGGTGCACACTGCCACACGAGTGGTCTGATCCACTGGAACGAACGAGAGATTCAT 329  
 |||||  
 Qy 384 LeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyr 403  
 |||||  
 Db 328 CTGAAGGACATTCAACACGATCGCAATGCAGTTCTGCGGAGGGAACAACCGAAGCTAC 269  
 |||||  
 Qy 404 PheGlyIlePheLysVal 409  
 |||||  
 Db 268 TTTGGCATCTTCAAGGTG 251

RESULT 4  
 CK135525/c  
 LOCUS

DEFINITION CK135525 532 bp mRNA linear EST 02-DEC-2003  
 RH51659.3prime RH Drosophila melanogaster normalized Head pFlc-1  
 Drosophila melanogaster cDNA clone RH51659 3 similar to CG12919:  
 FBan0012919 GO:() located on: 2R 46E1-46E1;: 08/05/2002, mRNA  
 sequence.

CK135525  
 CK135525.1 GI:38627461  
 EST.  
 Drosophila melanogaster (fruit fly)

ORGANISM  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

REFERENCE  
 1 (bases 1 to 532)  
 Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,  
 Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,  
 George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,  
 Misra,S., Mungall,C.J., Nunoo,J., Pacieb,J., Paragas,V., Park,S.,  
 Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and  
 Rubin,G.M.

TITLE  
 JOURNAL  
 COMMENT BDGP/HMI RH Drosophila EST Project  
 Unpublished (2001)  
 Other ESTs: RH51659.5prime  
 Contact: Stapleton, M.

BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
 Based upon the presence of a XhoI site followed by a run of 14 or  
 more T residues at the beginning of the sequence, this clone was  
 polyadenylated. The resulting Poly-T sequence has been removed. hit  
 genomic AE003831: arm:2R [4876906,5137815]  
 estimated-cyto:46B13-46E4: 08/05/2002  
 Plate: RH 516 row: E column: 11  
 High quality sequence stop: 382.

## FEATURES

source

1..532  
 Location/Qualifiers  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"  
 /clone="RH51659"  
 /sex="male and female"  
 /dev\_stage="Adult"  
 /lab\_host="DHS-alpha TonA"  
 /clone\_lib="RH Drosophila melanogaster normalized Head  
 pFlc-1"

/note="Organ: head; Vector: pFlc1; Site\_1: XhoI; Site\_2:  
 BamHI; Library was kindly generated by Piero Carninci at  
 the RIKEN. The library was normalized and excised using  
 Cre recombinase. Plasmid cDNA library."





```
/organism="Glossina morsitans morsitans"
/mol_type="mRNA"
/sub_species="morsitans"
/db_xref="taxon:37546"
/clone="Tse129f05.q1c"
/tissue_type="adult infected gut"
/clone_lib="Glossina morsitans morsitans adult infected
gut"
/notes="country: Zimbabwe; EST from adult gut infected with
T.brucei"

ORIGIN
Alignment Scores:
Pred. No.:      2.54e-08      Length:      539
Score:          206.50      Matches:      54
Percent Similarity: 42.00%      Conservative: 30
Best Local Similarity: 27.00%      Mismatches:  49
Query Match:    9.55%      Indels:      67
DB:             13      Gaps:        7

US-09-813-329-6 (1-409) x BX553466 (1-539)
QY      38  IleProLeuValLeuGlyPheIleGlyLeuGlyLeuValAlaIleLeuAlaLeuThr 57
Db      10  TTGGCAATTGTTTGACTTTTGTA-----CTCTGGGTGCTCG 48
QY      58  IleTrp-----GlnThrThrArg 63
Db      49  ATCTGGCATTAACTATTATACGNAATTAGACTCGCTAACGAAATAATACGCACAGT 108
QY      64  ValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsnLeuGln 83
Db      109  ATTAGAGATCTTCAAAGAGAGTCACTACTCAACAAGATCATTTGAAGTCTACAGAA 168
QY      84  ArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyrGluAsnAla 103
Db      169  CGATTGGGTATACCATATCTGGATGATTACATGATTTTGAAGAAGAGTATAAAATGTT 228
QY      104  Leu----- 104
Db      229  GTTGGAGTCTATTCCCGCAGCACTTACGATTAAGTTTAAAGATAAAGATCTTGACAAGAG 288
QY      105  -----IleAspTyrProLysLysVal-----AspGlyLeu 114
Db      289  AACGCAGAAAGTGAAGCGGAAGATTATGAGGAGACATCGTTTACAGGAGGACGATCTC 348
QY      115  ThrAspGluGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAsp 134
Db      349  AATGATAATGATAATGATGATGACGTCAT-----GATGACGATGAT 390
QY      135  AspAspValSerTyrSerValAspAspValGlyAlaAspTyrGluAspTyrThrAsp 154
Db      391  GATGATACCGATGACGGTAGCGAAGACGATGTTGAT---GACTATGAAGTATGATGAT 447
QY      155  MetLeuAsnLysLeuAsnAlaHisThrGlyThrThrProThrSerGluThrThrAla 174
Db      448  TTAATGGAGAATTCGTAATTATCAGAAAT----- 477
QY      175  GluGlyGluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAsp 194
Db      478  GAAGAGATAATAATAACGATGATAATGACTCGCGCGTCGTTAATGATAAATTTATATGAC 537

RESULT 7
AI253821
LOCUS      AI253821      526 bp      mRNA      linear      EST 03-DEC-1999
DEFINITION  AIMS-P.mon24 Giant tiger prawn eyestalk cDNA library Penaeus
            monodon cDNA 5', mRNA sequence.
ACCESSION  AI253821
VERSION    AI253821.1  GI:3858749
KEYWORDS   EST.
SOURCE     Penaeus monodon (black tiger shrimp)
ORGANISM   Penaeus monodon
            Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
```

```
Rumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
Penaeidae; Penaeus.
1 (bases 1 to 526)
Lehnert,S.A., Wilson,K.J., Byrne,K. and Moore,S.S.
Tissue-Specific Expressed Sequence Tags from the Black Tiger Shrimp
Penaeus monodon
Mar. Biotechnol. 1 (5), 465-476 (1999)
Contact: Wilson Kate J
Marine Biotechnology
Australian Institute of Marine Science
PMB 3, Townsville Mail Centre, Townsville QLD 4810, Australia
Tel: +61 7 4753 4444
Fax: +61 7 4772 5852
Email: kwilson@aims.gov.au
No significant database matches when searched using a filter to
remove low-complexity sequence, September 3 1998. When searched
with the filter turned off, shows homology to large proline-rich
protein bat2 from humans (p= 2.0e-07)
Insert Length: 1000 Std Error: 100.00
Seq primer: T3.

FEATURES             Location/Qualifiers
     source            1..526
     organism="Penaeus monodon"
     mol_type="mRNA"
     db_xref="taxon:6687"
     sex="Female"
     tissue_type="Eyestalk"
     dev_stage="Mature broodstock"
     clone_lib="Giant tiger prawn eyestalk cDNA library"
     notes="Vector: Lambda ZAP Express; Site 1: EcoRI; Site 2:
XhoI; mRNA isolated from total eyestalk tissue and used to
construct directional cDNA library in Lambda ZAP Express.
Clones excised in vivo to create phagemids in pBK-CMV for
sequencing."

ORIGIN
Alignment Scores:
Pred. No.:      3.55e-07      Length:      526
Score:          193.50      Matches:      59
Percent Similarity: 46.96%      Conservative: 26
Best Local Similarity: 32.60%      Mismatches:  70
Query Match:    8.95%      Indels:      26
DB:             9      Gaps:        8

US-09-813-329-6 (1-409) x AI253821 (1-526)
QY      231  GlnGluLysSerSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArg 250
Db      45  CAAGAAGATCCCGAGGAGGAAGACAGAGTCA-----ACGCCGCGCGC-----CACCAA 92
QY      251  ArgArgMetHisSerArgHisArgHisLeuLeuValArgLysAlaArgSerGluAspSer 270
Db      93  CCCCGT-----GGTCACCGTCGCCCACTTCGTGCCCCACGCCCCCGAACAGGAC--- 140
QY      271  ArgProAlaAlaHisPheHisLeuSerSerArg-ArgArgHisGlnGly---SerMetGl 289
Db      141  -----CGCCACCACACGACGAGTAGCGAAGA 170
QY      289  yTyrHisGlyAspMetTyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisP 309
Db      171  CGTACACGCGGAGTGGTCTCTCTCGCGCTGGATGGACAAGTAGGTCTCAACCCCAAGTA 230
QY      309  eGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGl 329
Db      231  CTCTCTCAGAGGGCGGTGTGTACGGTCAAGGAATCTGGACTGTATTACCTCTATGCGCA 290
QY      329  nileCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrPr 349
Db      291  GGTGCTGTACAGCTGGAGCGTTCGCGACGTGGGTTCAGGTGCTCGACACATATCCC 350
QY      349  oPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysValHisThrCysHi 369
Db      351  CATCATGAGTGC-----ACGTGGCTCCCGCTCAGCCGCTCTCTCT-----TCTTGCCA 398
```

```

QY 369 sThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAs 369
Db 399 CACCGGGGGCCACCTACCTCTCCAGAACGCTGCGCTACACATCCGGGACCTGGACCA 458
QY 389 nAspArgAsnAlaValLeuArgGluGlyArgAsnAsnArgSerTyrPheGlyIlePheLysVa 409
Db 459 CCACATGACCGCGTGAAGATGAGGAGAAC-----AGCTTCTTCGGCCTCGTGAAGCT 512
QY 409 1 409
Db 513 G 513

RESULT 8
LOCUS CNS017WQ 1101 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN37D20 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108596
VERSION AL108596.1 GI:5628900
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.

FEATURES             Location/Qualifiers
     source            1..1101
     organisms="Drosophila melanogaster"
     /mol_type="genomic DNA"
     /db_xref="taxon:7227"
     /clone="BACN37D20"
     /clone_lib="DrosBAC"
     /plasmid="pBelOBAC11"
     /note="end : T7"

ORIGIN
Alignment Scores:      8.22e-05      Length:      1101
Pred. No.:            172.00      Matches:      35
Score:                 67.21%      Conservative: 6
Best Local Similarity: 57.38%      Mismatches:  20
Query Match:          7.96%        Indels:       0
DB:                   29           Gaps:         0

US-09-813-329-6 (1-409) x CNS017WQ (1-1101)

QY 56 LeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLys 75
Db 256 TTGCAGAWMTGGCAGAAWNRGCTGATCGCATCTGGCAAGAACTGAAACCTGAAV 315
QY 76 ArgValValAspAsnLeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGlu 95
Db 316 CSARTCGTCRATAATCTCCAGCAACGTTTGGGCATAAACWATCTCGACGAATACGACGAT 375
QY 96 PheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThr 115
Db 376 TTCCAAAARGGFTCAGCAATTCGGTAAATCACTTTCCGAACATCAATATWWTACTCGGC 435

```

```

QY 116 Asp 116
Db 436 GAT 438

RESULT 9
LOCUS CNS07DXT 917 bp DNA linear GSS 08-JUL-2001
DEFINITION T3 end of clone BD0AA016E09 of library BD0AA from strain CBS 94 of
Candida tropicalis, genomic survey sequence.
ACCESSION AL441063
VERSION AL441063.1 GI:12224469
KEYWORDS GSS.
SOURCE Candida tropicalis
ORGANISM Candida tropicalis
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Mitosporic Saccharomycetales; Candida.
REFERENCE 1 (bases 1 to 917)
AUTHORS Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
PUBMED 11152876
REFERENCE 2 (bases 1 to 917)
AUTHORS Blandin,G., Ozier-Kalogeropoulos,O., Wincker,P., Artiguenave,F. and
Dujon,B.
TITLE Genomic exploration of the hemiascomycetous yeasts: 16. Candida
tropicalis
JOURNAL FEBS Lett. 487 (1), 91-94 (2000)
MEDLINE 20584726
PUBMED 11152891
REFERENCE 3 (bases 1 to 917)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web :
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

FEATURES             Location/Qualifiers
     source            1..917
     organisms="Candida tropicalis"
     /mol_type="genomic DNA"
     /strain="CBS 94"
     /db_xref="taxon:5482"
     /clone="BD0AA016E09"
     /clone_lib="BD0AA"
     /note="end : T3"

misc_feature          <10..>383
     /note="similar to Saccharomyces cerevisiae ORF YKL088w [
similarly to C.tropicalis hal3 protein, to C-term. of
Sis2p and to hypothetical protein YOR054c ]
1 putative frameshift(s)"
     /evidence=not_experimental

ORIGIN
Alignment Scores:      9.62e-05      Length:      917
Pred. No.:            170.00      Matches:      62
Score:

```

Percent Similarity: 42.79%      Conservative: 30  
Best Local Similarity: 28.84%      Mismatches: 73  
Query Match: 7.86%      Indels: 50  
DB: 29      Gaps: 10

US-09-813-329-6 (1-409) x CNS07DXT (1-917)

```
Qy 40 LeuValLeuGlyPheileGlyLeuGlyLeuValValAlaLeuLeuAlaLeuThrIleTrp 59
Db 297 TTAGTTTGTGTGATATTGGAATGGA-----GGTATCGGAGATGG 338
Qy 60 GlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLeu-----Lys 75
Db 339 GCTGATATT---GTCGATATCTTGAAGAAGAAATCAATTCATTAAGCTGAACAAAG 395
Qy 76 ArgValValAspAsnLeuGlnArgLeuGlyLeuLeuLeuLeuLeuLeuLeuLeu 95
Db 396 AGTTATTAGATGAA---GAAGAAGATTG-----AATGAAGAGGAGGAAGATGAG 446
Qy 96 PheGlnLysGluTyrGluAsnAlaLeuileAspTyrProLysLysValAspGlyLeuThr 115
Db 447 GATGAAGAAGAGGTGATGACGGC-----GAAGAAGATGGTATGATCGG 488
Qy 116 AspGluGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAsp 135
Db 489 GACGATGATGATGATGACGACGACGATGATGATGATGATGATGATGATGATGAT 539
Qy 136 AspValSerTyrSerSerValAspValGlyAlaAspTyrGluAspTyrThrAspMet 155
Db 540 GAC-----GACGACGAAGATGATGATGATGATGATGATGATGATGATGATGAT 581
Qy 156 LeuAsnLysLeuAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGlu 175
Db 582 GATAACCAAAATACAAATGTGACATCGGT-----TCATCCAAAGGATTAGCTAAT 632
Qy 176 GlyGluGlyThrAspSerAlaSerSerAlaSerAspAspAsnValPheAspAsp 195
Db 633 GGTAATCTGAAGCACCAGCAATCTAAAGATGACGAGAATGATGAATGATATTGAR--- 689
Qy 196 PheThrSerTyrAsnAlaHisLysLysGlnGluArgLysSerArgSerIleAlaAsp 215
Db 690 -----ATTACAGAT 698
Qy 216 ValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlySerSer 235
Db 699 CAAGGCAAAAGTGAGTCAGAAAGTCACACAGAGATGAGATATCAAGATAGTTCATCA 758
Qy 236 AsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArg 250
Db 759 ATTACTAAGAAACGAACCTGTTTACCGGCACCACTCAACTCGTGT 803
```

## RESULT 10

BU373036  
LOCUS BU373036 709 bp mRNA linear EST 28-NOV-2002  
DEFINITION 603812126F1 CSEQCHN74 Gallus gallus cdna clone CHEST800c8 5', mRNA sequence.

ACCESSION BU373036

VERSION BU373036.1 GI:25881037

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

1 (bases 1 to 709)

REFERENCE Boardman,P.E., Sans-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
AUTHORS Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.  
TITLE A Comprehensive Collection of Chicken cDNAs

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE 22335534

PUBMED 12445392

CONTACT Simon Hubbard

DEPARTMENT Department of Biomolecular Sciences

## FEATURES

## source

1. 709  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="Compton Line 151"  
/db\_xref="taxon:9031"  
/clone="CHEST800c8"  
/sex="Female"  
/dev\_stage="adult"  
/lab\_host="DH108"  
/clone\_lib="CSEQCHN74"  
/notes="Organ: kidney + adrenal; Vector: pBluescript II  
KS(+); Site 1: EcoRI; Site 2: NotI; This normalized  
library was constructed from 1 million independent clones.  
cDNA synthesis was initiated using an oligo(dT) primer,  
using methylated C in the first strand synthesis reaction.  
Following this first strand reaction, double-stranded cDNA  
was blunted, ligated to NotI adapters, digested with  
EcoRI, size-selected, and cloned into the NotI and EcoRI  
compatible sites of a custom modified MCS of the  
pBluescript (KS+) vector. The library was normalized in 2  
rounds using conditions adapted from Soares et al., PNAS  
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
(1996): 791, except that a significantly longer  
reannealing hybridization was used."

## ORIGIN

Alignment Scores:  
Pred. No.: 0.00334      Length: 709  
Score: 151.00      Matches: 46  
Percent Similarity: 49.34%      Conservative: 29  
Best Local Similarity: 30.26%      Mismatches: 63  
Query Match: 6.98%      Indels: 14  
DB: 13      Gaps: 7

US-09-813-329-6 (1-409) x BU373036 (1-709)

```
Qy 264 LysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArg 283
Db 146 AAGCGCGCTCCGAGACACCCAGCAGCTGTG---GTCACCTTCAAGGC----- 193
Qy 284 HisGlnGlySer---MetGlyTyrHisGlyAspMetTyrIleGly-----AsnAspAsn 300
Db 194 ---CAAGGCTCAGCCATCCCAAGTGAAGATGATCTTTTCAGGTGGAGTCTCAATGACTGG 250
Qy 301 GluArg-----AsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThr 317
Db 251 TCTCGCATCATATGAACCCCAAGGTGTTTAAAGCTGCATGCCCGCAGTGGGAGCTGGAG 310
Qy 318 ValThrAsnThrGlyLeuTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAsp 337
Db 311 GTACTGGTGACGCGCACATATCTATCTATGATGAGTATACATATAAATCTTCACAGAT 370
Qy 338 GlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrVal 357
Db 371 TTTCGCCAGCTATGAGTGGTGGTGGATGAATAAACCTTTCTGCAATGCACCTCGAGATATT 430
Qy 358 ProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGlu 377
Db 431 GAGACCGCGCAAGACCAAC---TTCAATACCTGCTATACGCGCGGGGTCTCCCTCTCAA 487
Qy 378 ArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGlu 397
Db 488 GCCAGGCAAGAAATTGCTGTGAATAATGGTCCAGCTGACATCTCC-----ATCAACATG 541
Qy 398 GlyAsnAsnArgSerTyrPheGlyIlePheLeuVal 409
Db 398 GlyAsnAsnArgSerTyrPheGlyIlePheLeuVal 409
```

University of Manchester Institute of Science and Technology  
(UMIST)  
PO Box 88, Manchester, M60 1QD, UK  
Tel: 01612008930  
Fax: 01612360409  
Email: Simon.Hubbard@umist.ac.uk.

US-09-813-329-6 (1-409) x CD851068 (1-746)

Qy 94 AspGluPheGlnLysGluTyrGluAsnAlaLeuLeuIleAspTyrProLysLysValAspGly 113  
Db 171 GATGAAGTGTATGATGAAGAAATTGACGCCATATTGGAT-----TCCACGAACCT 221  
Qy 114 LeuThrAspGluGluAspAspAspGlyLeuAspSerIleAlaAspAspGlu 133  
Db 222 GCTCTGGATGTCGCGGTATGATGATACGATATCTGGACGAATCGCTGACGACGAC 281  
Qy 134 AspAspAspValSerTyrSerValAspValGlyAlaAspTyrGluAspTyrThr 153  
Db 282 CATGATGATTGGTCGGTAATGCTAGTGATGATGAGGAATGGAGTT----- 329  
Qy 154 AspMetLeuAsnLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrThr 173  
Db 330 -----CCGGCAGATGTTTCAGAAAGATGAT 353  
Qy 174 AlaGluGlyGluGlyGluThrAsp-----SerAlaSerSerAlaSerAsnAsp 189  
Db 354 ---GAGGGGGAGGGTGATGATGATGTCAGCATTCGTGTGATGAGAGACGGAAGTGATGAC 410  
Qy 190 ---AspAsnValPheAspAspPheThrSerTyrAsnAlaHisLysLysLysGlnGluArg 208  
Db 411 GAGGATGACGTGTTGAA-----GTTAAGCCCTCGAGGAAGG 449  
Qy 209 LysSerArgSerIleAlaAspValArg----- 217  
Db 450 AAGACCCGAGGAGGCTGGCGGTCACCGTTTGCAGTCTTGAGGAATATGACCATTTG 509  
Qy 218 -----AsnGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlu 232  
Db 510 ATGGATGAAATGATGATAATGAAATGCAAAATGTCAGGGCAGTAAAGAAAGAAATCT 569  
Qy 233 LysSerSerAsnGluAlaThrSerLysGlu 242  
Db 570 AAGAAGGCTAAAGTCCCAATCTAAAG 599  
  
BI514804 550 bp mRNA linear EST 08-APR-2002  
LOCUS BB160016A20A06.5 Bee Brain Normalized Library, BB16 Apis mellifera  
DEFINITION cDNA clone BB160016A20A06 5', mRNA sequence.  
BI514804  
VERSION BI514804.1 GI:15365178  
KEYWORDS EST.  
SOURCE Apis mellifera (honeybee)  
ORGANISM Apis mellifera  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
Apoidea; Apis.  
1 (bases 1 to 550)  
Whitfield,C.W., Band,M.R., Bonaldo,M.F., Kumar,C.G., Liu,L.,  
Pardinas,J., Robertson,H.M., Soares,B. and Robinson,G.E.  
Annotated expressed sequence tags and cDNA microarrays for studies  
of brain and behavior in the honey bee  
Genome Res. 12 (4), 555-566 (2002)  
21929762  
11932240  
Contact: Gene E. Robinson  
Department of Entomology  
University of Illinois  
505 S. Goodwin Ave., Urbana, IL 61801, USA  
Tel: 217 265 0309  
Fax: 217 244 3499  
Email: generobi@life.uiuc.edu  
This research was funded by the University of Illinois Critical  
Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation  
Award in Functional Genomics to G.E. Robinson and an NSF  
Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.  
PCR Primers  
FORWARD: TAATACGACTCACTATAGG  
BACKWARD: ATTAACCTCACTAAG  
Plate: BB160016A20 row: A column: 06

Seq primer: AGCGATAACAATTTCACAGGA  
High quality sequence stop: 550.  
Location/Qualifiers  
1..550  
/organism="Apis mellifera"  
/mol\_type="mRNA"  
/strain="mixed strains of European bees, predominantly  
A.m. ligustica"  
/db\_xref="taxon:7460"  
/clone="BB160016A20A06"  
/sex="female"  
/tissue\_type="brain"  
/dev\_stage="adult worker honey bee"  
/lab\_host="DH10B"  
/clone\_lib="Bee Brain Normalized Library, BB16"  
/notes="Organ: Brain; Vector: pT7T3-Pac; Site 1: EcoRI;  
Site 2: NotI; The BB16 library was contributed by the  
Soares laboratory and it was constructed and normalized  
as described by Bonaldo, M.F., Lennon, G. and Soares,  
M.B. (1996), Genome Research 6(9): 791-806. RNA was  
prepared from dissected brains of adult worker bees of  
various ages and various behavioral groups."

ALIGNMENT SCORES:  
Pred. No.: 0.0122 Length: 550  
Score: 143.00 Matches: 32  
Percent Similarity: 55.56% Conservative: 23  
Best Local Similarity: 32.32% Mismatches: 42  
Query Match: 6.61% Indels: 2  
DB: 12 Gaps: 1  
US-09-813-329-6 (1-409) x BI514804 (1-550)

Qy 311 ThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrValTyrAlaGlnIle 330  
Db 53 TCCAACGATGGAGTGTATCCCAATTATGAACACGAGATTGTATTAGTTATGACACATT 112  
Qy 331 CysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPhe 350  
Db 113 CATTACAAATGACATGATGATGAAATGGATTTCATCTTCAAGTGACAACTCACTATT 172  
Qy 351 LeuGlnCysLeuAsnThrValProThrAsnMetProHisLysValHisThrCysHisThr 370  
Db 173 CTTCAATGATGATCGACAATTCGGCCATCGCGCAACATTAGCCAACTGTTGTTCTCA 232  
Qy 371 SerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAsp 390  
Db 233 GCTCAAGTGACACTTCTTAGAAGAGGAGACGTTTGTAGTTTTCAGGAGGCTAGCTCAC 292  
Qy 391 ArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409  
Db 293 AGATTCTGCTATTTTCGATAAGGAGAAC-----AGTTCTTGTGATTAGTGAACCTT 343  
  
RESULT 14  
BU798281 724 bp mRNA linear EST 23-OCT-2003  
LOCUS BU798281  
DEFINITION SJF2ACB08 SJF Schistosoma japonicum cDNA, mRNA sequence.  
ACCESSION BU798281  
VERSION BU798281.1 GI:28355288  
KEYWORDS EST.  
SOURCE Schistosoma japonicum  
ORGANISM Schistosoma japonicum  
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.  
1 (bases 1 to 724)  
Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,  
Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,  
Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,  
McManus,D.P., Xue,C.L., Feng,Z., Chen,Z. and Han,Z.G.  
Evolutionary and biomedical implications of a Schistosoma japonicum  
complementary DNA resource  
Nat. Genet. 35 (2), 139-147 (2003)

```

MEDLINE      22879925
COMMENT      Contact: Zeguang Han
              Chinese National Human Genome Center at Shanghai
              351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
              201203, P. R. China
              Tel: 86-21-50801919(ex.45)
              Fax: 86-21-50801922
              Email: hanzg@chgc.sh.cn.

FEATURES     source
              1..724
              /organism="Schistosoma japonicum"
              /mol_type="mRNA"
              /db_xref="taxon:6182"
              /sex="female"
              /tissue_type="whole body"
              /dev_stage="adult"
              /lab_host="rabbits"
              /clone_lib="SJF"

ORIGIN
Alignment Scores:
Pred. No.:      0.0178      Length:      724
Score:          143.00      Matches:    45
Percent Similarity: 43.26%  Conservative: 32
Best Local Similarity: 25.28% Mismatches:  79
Query Match:    6.61%      Indels:     22
DB:             13         Gaps:       5

US-09-813-329-6 (1-409) x BU798281 (1-724)
QY   91 AspGluPheAspGluPhe-----GlnLysGluTyrGluAsnAlaLeuLeuAspTyr 107
    |||
Db   124 GATGATTTGATGAGTTCGGTCGTAAGAAAGAAAGTATCGCGAAGTCAAGTGAACAA 193
    |||

QY   108 ProLysLysValAspGlyLeu-----
    |||
Db   184 CTTAGTACCATGACTCGTCGATCTCGTAACAATAATGCTACACCAGGACATGATCAT 243
    |||

QY   115 ThrAspGluGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAsp 134
    |||
Db   244 AATGATGAATGTCATTAACAAGGATGATGATGATGATGATGATGATGATGATGATGAT 303
    |||

QY   135 AspAspValSerTyrSerSerValAspValGlyAlaAspTyrGluAspTyrThrAsp 154
    |||
Db   304 GATGATGCTGATCTCTTAAGTATGATATTGGGGGGCAGAAATGACGATTTAGGTGAC 363
    |||

QY   155 MetLeuAsnLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrAla 174
    |||
Db   364 -----AAGTTATCAAAATGCACAA-----AATCTCGTATCTTCTCAAGTATCTTACC 408
    |||

QY   175 GluGlyGluGlyGluThrAspSerAlaSerAlaSerAlaSerAspAspAspAspValPheAsp 194
    |||
Db   409 AATAAAAAATCTCGAAGTCGTAGTCCATGCTCTTCCTCGGACTCAGACTCAGATTCTTCC 468
    |||

QY   195 AspPheThrSerTyrAsnAlaHisLysLysLysGlnGluArgLysSerArgSerIleAla 214
    |||
Db   469 TCATCATCCAGCCAGACTTCACATTCCTCAACATCAGATTCATCATCAAGTTCTCGTAGT 528
    |||

QY   215 AspValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSer 234
    |||
Db   529 GAGTCATCTTCCAAACCCGAAAAAT---CAATCCAAACCGGCAAGAGACCTCATCTCTAAAAAT 585
    |||

QY   235 SerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgArg 252
    |||
Db   586 GGCAATACGGGCAACGATAGTGACAGTTCACCTACCGTAATACCAAGCGAAAA 639
    |||

RESULT 15
BU803616/c
LOCUS      BU803616
DEFINITION SJFAWE05 SJF Schistosoma japonicum cDNA, mRNA sequence.
ACCESSION  BU803616
VERSION    BU803616.1 GI:28360623
KEYWORDS   EST.

SOURCE      Schistosoma japonicum
ORGANISM    Schistosoma japonicum
REFERENCE   1 (bases 1 to 732)
AUTHORS     Hu, W., Yan, Q., Shen, D.K., Liu, F., Zhu, Z.D., Song, H.D., Xu, X.R.,
              Wang, Z.J., Rong, Y.P., Zeng, L.C., Wu, J., Zhang, X., Wang, J.J.,
              Xu, X.N., Wang, S.Y., Fu, G., Zhang, X.L., Wang, Z.Q., Brindley, P.J.,
              McManus, D.P., Xue, C.L., Feng, Z., Chen, Z. and Han, Z.G.
              Evolutionary and biomedical implications of a Schistosoma japonicum
              complementary DNA resource
              Nat. Genet. 35 (2), 139-147 (2003)
              22879925
              Contact: Zeguang Han
              Chinese National Human Genome Center at Shanghai
              351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
              201203, P. R. China
              Tel: 86-21-50801919(ex.45)
              Fax: 86-21-50801922
              Email: hanzg@chgc.sh.cn.
              Location/Qualifiers
                  1..732
                  /organism="Schistosoma japonicum"
                  /mol_type="mRNA"
                  /db_xref="taxon:6182"
                  /sex="female"
                  /tissue_type="whole body"
                  /dev_stage="adult"
                  /lab_host="rabbits"
                  /clone_lib="SJF"

ORIGIN
Alignment Scores:
Pred. No.:      0.0181      Length:      732
Score:          143.00      Matches:    45
Percent Similarity: 43.26%  Conservative: 32
Best Local Similarity: 25.28% Mismatches:  79
Query Match:    6.61%      Indels:     22
DB:             13         Gaps:       5

US-09-813-329-6 (1-409) x BU803616 (1-732)
QY   91 AspGluPheAspGluPhe-----GlnLysGluTyrGluAsnAlaLeuLeuAspTyr 107
    |||
Db   593 GATGATTTGATGAGTTCGGTCGTAAGAAAGAAAGTATCGCGAAGTCAAGTGAACAA 534
    |||

QY   108 ProLysLysValAspGlyLeu-----
    |||
Db   533 CTTAGTACCATGACTCGTCGATCTCGTAACAATAATGCTACACCAGGACATGATCAT 474
    |||

QY   115 ThrAspGluGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAsp 134
    |||
Db   473 AATGATGAATAATGTCATTAACAAGGATGATGATGATGATGATGATGATGATGATGAT 414
    |||

QY   135 AspAspValSerTyrSerSerValAspValGlyAlaAspTyrGluAspTyrThrAsp 154
    |||
Db   413 GATGATGCTGATCTCTTAAGTATGATATTGGGGGGCAGAAATGACGATTTAGGTGAC 354
    |||

QY   155 MetLeuAsnLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrAla 174
    |||
Db   353 -----AAGTTATCAAAATGCACAA-----AATCTCGTATCTTCTCAAGTATCTTACC 309
    |||

QY   175 GluGlyGluGlyGluThrAspSerAlaSerAlaSerAlaSerAspAspAspAspValPheAsp 194
    |||
Db   308 AATAAAAAATCTCGAAGTCGTAGTCCATGCTCTTCCTCGGACTCAGACTCAGATTCTTCC 249
    |||

QY   195 AspPheThrSerTyrAsnAlaHisLysLysLysGlnGluArgLysSerArgSerIleAla 214
    |||
Db   248 TCATCATCCAGCCAGACTTCACATTCCTCAACATCAGATTCATCATCAAGTTCTCGTAGT 189
    |||

QY   215 AspValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSer 234
    |||
Db   188 GAGTCATCTTCCAAACCCGAAAAAT---CAATCCAAACCGGCAAGAGACCTCATCTCTAAAAAT 132
    |||
```

Qy 235 SerAsnGluAlaThrSerLySGluSerProAlaProLeuHisHisArgArgArg 252  
Db 131 GGCAATACGGGCAACGATAGTGACAGCTCACCTACCGTAAATCACAGCGAAA 78

Search completed: September 15, 2004, 11:25:28  
Job time : 4198 secs



